



universidad
de león

UNIVERSITY OF LEÓN

DEPARTMENT OF ELECTRICAL, SYSTEMS AND AUTOMATIC
ENGINEERING

WEAR CHARACTERIZATION OF THE CUTTING TOOL IN MILLING PROCESSES USING SHAPE AND TEXTURE DESCRIPTORS

A dissertation supervised by

PROF. DR. ENRIQUE ALEGRE GUTIÉRREZ,

PROF. DR. ROCÍO ALAIZ RODRÍGUEZ

and submitted by

MARÍA TERESA GARCÍA ORDÁS

in fulfillment of the requirements for the Degree of

PHILOSOPHIÆDOCTOR (PH.D.)

July 2017



universidad
de león

UNIVERSIDAD DE LEÓN

DEPARTAMENTO DE INGENIERÍA ELÉCTRICA Y DE SISTEMAS Y
AUTOMÁTICA

CARACTERIZACIÓN DEL DESGASTE DE LA HERRAMIENTA DE CORTE EN PROCESOS DE FRESADO UTILIZANDO DESCRIPTORES DE FORMA Y TEXTURA

Tesis doctoral dirigida por

EL PROF. DR. ENRIQUE ALEGRE GUTIÉRREZ,

LA PROF. DR. ROCÍO ALAIZ RODRÍGUEZ

y desarrollada por

MARÍA TERESA GARCÍA ORDÁS

a fin de optar al grado de

DOCTOR POR LA UNIVERSIDAD DE LEÓN

León, julio de 2017

Abstract

This thesis proposes several methods to evaluate the wear level of tools in milling processes. Using shape descriptors based on contour, moments and orientations, and also texture ones, the wear regions of the cutting inserts were characterised and using different classifiers the wear was modelled. The proposed models were evaluated on classic shape datasets as well as in new insert datasets created specifically for this problem.

The motivation behind our work was that optimizing tool replacement operations may produce a significant improvement in manufacturing efficiency and competitiveness. It is known that changing tools at the right moment is essential, not only because of the cost of cutting tools themselves, but also for the indirect costs due to the unproductive time needed to carry out the tool replacement. This thesis addresses the task of tool wear monitoring in edge profile milling processes using computer vision and machine learning techniques allowing to replace the tools at the optimum moment and thus saving direct and indirect costs.

We have created three insert datasets to assess our solutions in similar conditions to the real environment: Insert edges dataset, Insert region dataset, and Insert High-Resolution Dataset.

The general shape can be extract from the boundary pixels of an image. For this reason, our first proposal for shape description was a contour-based approach. We proposed RCPDH and CPDH36R, both based on CPDH (Contour points distribution histogram).

Next, we decided to describe the images using both moments and orientations. We introduced two new methods named aZIBO and B-ORCHIZ, both of them based on ZMEG approach. They are composed for global and local shape descriptors. The combination of local and global descriptors allows us to have an idea of the general form and to access the local details of the form. In this way, we take advantage of both methods of description.

B-ORCHIZ proved to be the best moment-orientation descriptor obtaining an accuracy up to 91.92% in the shape description problem.

Results were promising not only for generic shape datasets but also in real

applications obtaining results up to 87.02% in terms of accuracy in the tool wear monitoring problem.

As B-ORCHIZ descriptor proved to be useful in the tool wear monitoring field, we explored the combination of it with our shape descriptor proposal, ShapeFeat. ShapeFeat takes into account ten different features extracted from the binary region of each image.

Combining ShapeFeat with B-ORCHIZ and evaluating three types of fusion methods, we obtained results which improves the performance significantly, yielding a 91.44% of hit rate in the complete dataset evaluation using two classes and a 82.90% using three classes.

Our final proposal consisted of a new online, low cost and fast approach to determine the state and categorize the wear of cutting tools used in edge profile milling processes.

The method proposed is based on dividing the region of the cutting edge in patches. We presented five alternatives to carry out this division and we described each patch with different texture methods. The experiments were carried out using a Support Vector Machine (SVM) with intersection kernel. Results shown that our proposal "Small Edge Division" (SED) outperforms the other region division proposals achieving an Fscore value of 0.9032.

Resumen

En esta tesis se proponen múltiples métodos para evaluar el nivel de desgaste de herramientas en procesos de fresado. Hemos caracterizado la región de desgaste de las herramientas de corte usando descriptores basados en forma, contorno, momentos y orientaciones y también en textura. Además, se utilizaron diferentes clasificadores para modelar el desgaste. Los modelos propuestos fueron evaluados a su vez utilizando conjuntos de imágenes de forma clásicos y también conjuntos de imágenes de plaquitas creados específicamente para este problema.

La motivación que hay detrás de este trabajo, es optimizar las operaciones de reemplazamiento para poder mejorar la competitividad de las empresas.

Las técnicas conservadoras utilizadas hasta la fecha hacen que las herramientas sean reemplazadas mucho antes de que sea necesario, por el gran coste que conlleva su rotura mientras se está mecanizando, de manera que todavía les quedaría vida útil. Sin embargo, reemplazar las herramientas en el momento correcto es esencial, no solo por el coste derivado de las mismas, si no también por los costes indirectos ocasionados por el tiempo que se necesita para llevar a cabo la sustitución.

Esta tesis aborda la tarea de monitorizar el desgaste de las herramientas de corte en procesos de mecanizado utilizando técnicas de visión artificial y técnicas de aprendizaje máquina.

Para llevar a cabo este objetivo, se crearon tres conjuntos de imágenes de plaquitas de corte: conjunto de imágenes de filo de corte, conjunto de imágenes de regiones de plaquitas de corte, y el conjunto de imágenes de plaquitas en alta resolución.

La forma general de una imagen se puede extraer a partir de los puntos de contorno. Por esta razón, nuestra primera alternativa para la descripción de formas fue una propuesta basada en la descripción del contorno. Propusimos RCPDH y CPDH36R, ambos basados en CPDH (Contour Points Distribution Histogram).

Después de la descripción basada en contorno, decidimos utilizar métodos de descripción basados en momentos y orientaciones. En este sentido introducimos dos nuevos métodos llamados aZIBO y B-ORCHIZ, ambos basados en el método ZMEG. Están formados por descriptores de forma globales y locales. Los descriptores

locales se usan para caracterizar pequeños parches de la imagen, mientras que los descriptores globales nos dan información sobre la imagen global en conjunto. Sin embargo, no son tan robustos como los descriptores locales porque no son sensibles a cambios en partes pequeñas o concretas de la imagen.

B-ORCHIZ resultó ser la mejor de nuestras propuestas basadas en momentos y orientaciones y obtuvo una precisión del 91,92 % en el problema de descripción de forma .

Los resultados son prometedores no solo en los conjuntos de imágenes de formas genéricas, si no también en nuestra aplicación real de monitorizar el desgaste de la herramienta.

Como el descriptor B-ORCHIZ demostró ser útil en el campo de monitorización del desgaste de la herramienta, exploramos su combinación con nuestra propuesta de descriptor de forma, ShapeFeat. ShapeFeat tiene en cuenta diez características diferentes extraídas de la región binaria de cada imagen. Combinando ShapeFeat con B-ORCHIZ por medio de tres técnicas diferentes de fusión, se obtienen resultados que mejoran las tasas de acierto significativamente obteniendo un 91.44 % en el conjunto de imágenes de regiones de plaquitas completo, teniendo en cuenta dos clases y 82.90 % para la clasificación en tres clases de desgaste.

Nuestra propuesta final se trata de un método nuevo, en línea, de bajo coste y rápido para determinar el estado y categorizar el desgaste de herramientas de corte usadas en procesos de fresado. El método propuesto está basado en la división de la región de desgaste del filo de corte en parches para su descripción individual. Propusimos cinco alternativas de división y se describió cada parche utilizando descriptores de textura. Los experimentos se llevaron a cabo usando como clasificador, máquinas de vector soporte (SVM) con kernel intersección. Los resultados muestran que nuestra propuesta de división llamada "Small Edge Division" (SED) mejora a las otras cuatro propuestas y obtiene valores de Fscore de 0.9032.

Contents

List of Figures	IV
List of Tables	X
Acknowledgements	XV
1. Introduction	1
1.1. Motivation	1
1.2. Objectives	2
1.3. Main contributions	3
1.4. Thesis Organization	4
2. State of the art	7
2.1. Tool Wear monitoring	7
2.1.1. Indirect methods	7
2.1.2. Direct methods	11
2.2. Shape based descriptors	11
2.2.1. Shape parameters	12
2.2.2. One dimensional function methods	13
2.2.3. Polygonal approximation	15
2.2.4. Moments	16
2.2.5. Shape transform domains	18
2.2.6. Shape based description in tool wear monitoring	21
2.3. Texture based descriptors	22
2.3.1. Gray Level Cooccurrence Matrices	22
2.3.2. Local Binary Patterns	23
2.3.3. Texture based description in tool wear monitoring	23
3. Datasets	25
3.1. Shape retrieval datasets	25
3.2. Insert edges dataset	26
3.2.1. Cropping of the cutting edges	28

3.3. Insert regions datasets	30
3.4. Insert High resolution dataset for texture description	31
4. Contour based descriptors	35
4.1. Methodology	35
4.1.1. CPDH	35
4.1.2. CPDH36R	36
4.1.3. RCPDH	38
4.2. Shape recognition	39
4.2.1. Experimental setup	39
4.2.2. Results	41
4.3. Conclusions	42
5. Moments and orientations based descriptors	45
5.1. Methodology	46
5.1.1. ZMEG	46
5.1.2. aZIBO	47
5.1.3. B-ORCHIZ	48
5.2. Shape recognition application	50
5.2.1. Experimental setup	50
5.2.2. Results	50
5.3. Tool Wear monitoring	52
5.3.1. Experimental setup	52
5.3.2. Results on the insert edge dataset	52
5.3.3. Results on the insert region dataset	53
5.4. Conclusions	62
6. Fusion of region and moment descriptor	65
6.1. Methodology	65
6.1.1. Shape descriptor: ShapeFeat	65
6.2. Combination of Shape and Contour descriptors	67
6.2.1. Early fusion	69
6.2.2. Intermediate fusion: Co-transduction method	69
6.2.3. Late fusion	70
6.3. Tool Wear monitoring	70
6.3.1. Experimental setup	70
6.3.2. Insert Region dataset results	71
6.4. Conclusions	74
7. Patches based descriptor	75
7.1. Methodology	76
7.1.1. Regions configuration	78

CONTENTS

7.1.2. Texture descriptors	79
7.2. Tool Wear monitoring	82
7.2.1. Experimental setup	82
7.2.2. Threshold selection	83
7.2.3. Region wear configuration and description evaluation	84
7.3. Conclusions	86
8. Conclusions and outlook	87
8.1. Work summary	87
8.2. General contributions	88
8.3. Outlook	90
9. Conclusiones y perspectiva	93
9.1. Resumen del trabajo	93
9.2. Contribuciones generales	94
9.3. Perspectiva	97
Bibliography	99
Annex A: Research Activities	107
Annex B: Summary of the dissertation in Spanish	

List of Figures

1.1.	Left: overview of the machine. Right: detail of the head.	2
1.2.	Example of a damaged insert.	2
2.1.	Cutting force components on a single point during the turning process.	9
2.2.	Illustration of convex hull	12
2.3.	(a) Contours normalized to 128 points. The star indicate the starting point. (b) curvature function computed in clockwise.	13
2.4.	(a) Original contour. (b) the area function of (a).	14
2.5.	A few stages of polygon evolution according to a relevant measure.	15
2.6.	Splitting method for polygonal approximation.	16
3.1.	a) Kimia25 database composed by 25 samples divided in 6 classes. b) Kimia99 database composed by 9 classes with 11 samples per class.	25
3.2.	Example of 4 classes with 20 samples of MPEG2 dataset.	26
3.3.	Insert dataset illustration. Inserts on the left present high wear while inserts on the right present low wear taking into account the left cutting edge.	26
3.4.	Genie M1280 (left side), lateral view (center image), back view (right side).	27
3.5.	Image capture prototype. The support in which the camera is placed and the LEDs bars employed are shown.	27
3.6.	(a) Masked image of an insert. (b) Binary image after applying a threshold of 0.01. (c) Image with the center circle removed. (d) Image after sobel filter convolution. (e) Image after applying morphological operations (dilatation, opening). (f) Vertical projection of (e) to identify the edges.	28
3.7.	Rotated crops example.	29
3.8.	This is an example of the obtained dataset for the images shown in Figure 3.3. Each of the previous insert gray-scale images with masked background follow a preprocessing step that yields four new images, one for each cutting edge in horizontal position.	30

3.9. (a) The image containing both types of edges: The complete edge placed in horizontal position and the imcompletes edges in vertical position. (b) The complete edge and (c) Incomplete edges from vertical edges of (a). 30

3.10. Example of the captured images. 32

3.11. Cutting edge images. In the first row, intact edges are shown. In the second one, edges with different kind of wear are displayed. 33

4.1. Contour points extraction using Belongie method. 36

4.2. Circumference circumscribed to the image contour (left) and circle partition in 36 bins ($v=12, u=3$) (right). 37

4.3. First step to construct CPDH36R: 36SEQ extraction. 38

4.4. CPDH36R constructed by multiplying the number of points of each bin by its corresponding normalized radius 38

4.5. RCPDH descriptor. 39

4.6. kNN mean distance example taking $k=7$ 40

4.7. Hit rate of the different kNN methods applied on the CPDH36R descriptor using Kimia25 dataset (left) and Kimia99 dataset (right). 42

4.8. Hit rate of the different kNN methods applied on the CPDH36R descriptor using MPEG2 (left) and MPEG7 (right) datasets. 43

5.1. Steps followed to obtaining the eight quantized orientations. 46

5.2. Example of the orientations taken into account for the boundary point shown in red for the composition of the EGCM (left). Example of the composition of the edge gradient co-occurrence matrix taking into account the boundary point shown in this figure (right). 47

5.3. Example of the invariance to rotation of the IEGCM. Red dotted line represents the orientation between the furthest points in the boundary. The red arrow corresponds to the gradient orientation of the point whose IEGCM is calculated. The green arrows are taken into account (because they are located inside the neighborhood) whereas the blue arrows represent points that are not in the influence area of the selected one. 48

5.4. Example of the Boundary Orientations Chain (BOC) rotation invariance. The red arrow represents the maximal distance between points in the contour of the image. The green arrow is the orientation of the first point of the chain. The values assigned to each orientation are shifted locating the green arrow in the top of the diagram (value 0). 49

5.5. Precision-Recall curves of the classification using Kimia-99, MPEG-2 and MPEG-7 dataset with B-ORCHIZ, aZIBO and ZMEG. 51

5.6. kNN classification for the six distances and different values of k using aZIBO, Hu and Flusser and two kinds of labeling datasets: L-M-H and L-H classes.	54
5.7. Examples of misclassified images: (a) low wear inserts misclassified as high wear and (b) high wear inserts missclassified as low wear. . .	56
5.8. Precision-Recall curves (L-H classification) using B-ORCHIZ, aZIBO and ZMEG for Insert-C, Insert-I and the whole Insert dataset (from left to right).	57
5.9. Classification results using B-ORCHIZ descriptor and the wear area size. In (a) the results for L-H experiment. In (b) the results for L-M-H experiment.	58
5.10. Dendrogram of the clustering obtained using B-ORCHIZ (left), aZIBO (center) and ZMEG(right) descriptors for the Insert dataset and the Insert-C and Insert-I subsets.	59
5.11. Representative images for three levels of wear (L: first row, M: second row and H: third row) for the Insert-C subset using clustering. Insert wear described with B-ORCHIZ, aZIBO and ZMEG with its associated ground truth (columns from left to right) for automatic clustering. Images are aligned and the average value of each pixel of the image representing the gray level of the representative image is calculated.	61
5.12. Representative images for (a) the Insert-C subset and (b) the Insert-I subset, L-H labelling (from left to right).	62
5.13. Representative images for (a) the Insert-C subset and (b) the Insert-I subset, L-M-H labelling (from left to right). First row represents the surface and second row represents the contour for the plan view. . .	62
6.1. Ellipse with centre O . The foci coordinates are $F'(-c, 0)$ and $F(0, c)$.	66
6.2. Two examples of the ShapeFeat features extraction. With the information shown in the first column, we can extract the eccentricity and R . With the information shown in the second one, we can extract the extent and finally, the smallest convex polygon shown in the third column, is necessary to extract the Convex Area and the Solidity. . .	68
6.3. (a) Early fusion of the contour and shape descriptors.(b) Intermediate fusion using co-transduction method. (c) Late fusion using Bayes average.	68
6.4. Results for binary (a) and ternary (b) classification using fusion methodology with B-ORCHIZ and ShapeFeat. Furthermore, results achieved by the descriptors by themselves are represented as straight lines over the graphs.	73

7.1. Schema of our proposed system for tool wear monitoring using computer vision.	77
7.2. All the different configurations of the patches over the cutting edges.	78
7.3. Local Binary Pattern process over one pixel in gray scale level using a neighbourhood of radius 1 and 8 neighbours.	80
7.4. Recall values for all wear region configuration methods using different descriptors and threshold values.	84
7.5. Results of accuracy, precision, recall and FScore with all the methods and descriptors for the selected threshold value equal to one.	85
1. a) Conjunto de imágenes Kimia25 compuesto por 25 muestras repartidas en 6 clases. b) Conjunto de imágenes Kimia99 compuesto por 9 clases con 11 muestras por clase.	16
2. Ejemplo de 4 clases con 20 muestras del conjunto de imágenes MPEG2.	16
3. Ejemplo del conjunto de imágenes. Las plaquitas de la izquierda presentan alto desgaste mientras que las de la derecha presentan bajo desgaste teniendo siempre en cuenta el filo izquierdo de corte. . . .	17
4. Ejemplo del conjunto de imágenes del filo de corte.	18
5. (a) Imagen que contiene los dos tipos de filos: Filo completo, posicionado en horizontal y filos incompletos en posición vertical. (b) Filo completo (c) Filos incompletos que representan los filos verticales en (a).	19
6. Example of the captured images.	20
7. imágenes del filo de corte. En la primera fila se muestran los filos intactos. En la segunda, filos con diferentes niveles de desgaste.. . .	21
8. Extracción de los puntos de contorno usando el método de Belongie	22
9. Circunferencia circunscrita al contorno de la imagen (izquierda) y círculo particionado en 36 porciones ($v=12$, $u=3$) (derecha).	23
10. CPDH36R construido multiplicando el número de puntos de cada porción por su correspondiente radio normalizado	24
11. Primer paso para construir RCPDH: extracción de CPDH36SEQ. . . .	24
12. Descriptor RCPDH.	25
13. Tasa de acierto para los diferentes métodos kNN aplicados en el descriptor CPDH36R usando los conjuntos Kimia25 (izquierda) y Kimia99 (derecha).	27
14. Tasa de acierto para los diferentes métodos kNN aplicados en el descriptor CPDH36R usando los conjuntos MPEG2 (izquierda) y MPEG7 (derecha).	27
15. Pasos seguidos hasta obtener las ocho orientaciones cuantizadas. . .	29

16.	Ejemplo de las orientaciones que se tienen en cuenta para un punto de contorno, mostrado en rojo, para la construcción de EGCM (izquierda). Ejemplo de la construcción de EGCM teniendo en cuenta el punto mostrado en la figura (derecha).	29
17.	Ejemplo de invarianza a la rotación de IEGCM. La línea roja punteada representa la orientación entre los puntos más alejados del contorno. La línea roja se corresponde con la orientación del gradiente del punto para el que se está calculando el IEGCM. Las flechas verdes son los puntos que se tienen en cuenta (porque están dentro del vecindario) mientras que las azules representan puntos fuera del área de influencia del punto que estamos tratando.	30
18.	Ejemplo de construcción de la cadena de orientaciones del borde (BOC) invariante a la rotación. La flecha roja representa la distancia entre los puntos más alejados del contorno. La verde es la orientación del primer punto de la cadena. Los valores se asignan a cada orientación secuencialmente empezando por la flecha verde, la cual colocamos en la parte alta del diagrama como punto de origen (valor 0).	31
19.	Curvas PR para los conjuntos de imágenes Kimia-99, MPEG-2 y MPEG-7 descritos con B-ORCHIZ, aZIBO y ZMEG.	33
20.	Clasificación con kNN usando seis distancias y diferentes valores de k , para aZIBO, Hu y Flusser y dos categorías de etiquetado del conjunto de imágenes: L-M-H y L-H (3 clases de desgaste y dos clases de desgaste)	35
21.	Curvas PR describiendo con B-ORCHIZ, aZIBO y ZMEG los conjuntos de imágenes Insert-C, Insert-I e Insert (de izquierda a derecha).	37
22.	Dendrogramas de agrupamientos obtenidos usando los descriptores B-ORCHIZ (izquierda), aZIBO (centro) y ZMEG(derecha) para el conjunto de imágenes Insert, Insert-C e Insert-I.	38
23.	Elipse con centro O . Las coordenadas del foco son $F'(-c, 0)$ y $F(0, c)$	41
24.	Dos ejemplos de extracción de características de ShapeFeat. Con la información mostrada en la primera columna podemos extraer la excentricidad y el parámetro R y con la mostrada en la segunda, podemos extraer la extensión (extent). Finalmente, el polígono convexo más pequeño mostrado en la tercera columna, es necesario para extraer el área convexa (ConvexArea) y la solidez (Solidity). . .	43
25.	(a) Fusión temprana para descriptores de forma y contorno. (b) Fusión intermedia usando el método de Co-transduction. (c) Fusión tardía usando la media de Bayes.	44

26.	Resultados para la clasificación binaria (a) y ternaria (b) usando la metodología de fusión con B-ORCHIZ y ShapeFeat. Además, se muestra también los resultados alcanzados por cada descriptor por separado, representados como líneas rectas en la gráfica.	47
27.	Todas las diferentes configuraciones de división del filo	49
28.	Proceso de extracción de LBP para un pixel en escala de grises usando un vecindario de radio 1 y 8 orientaciones diferentes.	51
29.	Recall para todas las configuraciones de la zona de desgaste usando diferentes descriptores y diferentes umbrales.	55
30.	Resultados de accuracy, precision, recall y F-Score de todos los métodos para todas las configuraciones de regiones y con umbral=1.	56

List of Tables

3.1. Genie M1280 specifications.	27
3.2. Mean area of the wear for the L-H labelling and their standard deviation (mm^2)	31
3.3. Mean area of the wear for the L-M-H labelling and their standard deviation (mm^2)	31
4.1. Kimia25 and Kimia99 classification using CPDH original RCPDH and CPDH36R methods with k=1.	41
4.2. MPEG7 and MPEG2 classification using CPDH original and CPDH36R methods with k=1.	42
5.1. Classification accuracy in % of our proposal B-ORCHIZ against aZIBO and ZMEG using kNN (with different distances) in the Kimia-99 (first), MPEG-2 (second) and MPEG-7 (third) database.	51
5.2. Inserts classification using aZIBO against classical Hu and Flusser descriptors with k equals to nine and using Cosine, ChiSquare, SqDist and Cityblock distances with 3 classes labeling.	53
5.3. Inserts classification using aZIBO against classical Hu and Flusser descriptors with k equals to nine and using Cosine, ChiSquare, SqDist and Cityblock distances with 2 classes labeling.	53
5.4. Classification Accuracy (in %) using SVM (linear, quadratic and intersection kernels) of the two and three-class scenarios for the Insert-C dataset.	55
5.5. Classification Accuracy (in %) using SVM (linear, quadratic and intersection kernels) of the two and three-class scenarios for the Insert-I dataset.	55
5.6. Classification Accuracy (in %) using SVM (linear, quadratic and intersection kernels) of the two and three-class for the Insert dataset (full dataset).	56

6.1. Classification accuracy in % of ZMEG, aZIBO, B-ORCHIZ and ShapeFeat using SVM with Intersection kernel for the complete, Insert-C, Insert-I dataset (from left to right) for two and three wear levels.	71
6.2. Classification accuracy in % of the combination of ShapeFeat with BORCHIZ and other descriptors like BCF (Bag of contour fragments), Hog (Histogram of oriented gradients) and SC (Shape Context) . . .	74
1. Clasificación Kimia25 y Kimia99 usando CPDH original RCPDH y CPDH36R con k=1.	26
2. Clasificación MPEG7 y MPEG2 usando CPDH original yCPDH36R con k=1.	26
3. Tasas de acierto en % de nuestra propuesta B-ORCHIZ contra aZIBO y ZMEG usando kNN (con diferentes distancias) en los conjuntos de imágenes Kimia-99 (primero), MPEG-2 (segundo) y MPEG-7 (tercero).	33
4. Clasificación de las plaquitas usando nuestra propuesta aZIBO y dos descriptores clásicos: Hu y Flusser con valor de k igual a nueve y distancias coseno, chi cuadrado, sqDist y cityblock y etiquetado en tres clases diferentes de desgaste.	34
5. IClasificación de las plaquitas usando nuestra propuesta aZIBO y dos descriptores clásicos: Hu y Flusser con valor de k igual a nueve y distancias coseno, chi cuadrado, sqDist y cityblock y etiquetado en dos clases diferentes de desgaste.	34
6. Resultados de precisión (en %) usando SVM (con kernels lineal, cuadrático e intersección) para las clasificaciones binaria y ternaria y el conjunto de imágenes the Insert-C.	36
7. Resultados de precisión (en %) usando SVM (con kernels lineal, cuadrático e intersección) para las clasificaciones binaria y ternaria y el conjunto de imágenes the Insert-I.	36
8. Resultados de precisión (en %) usando SVM (con kernels lineal, cuadrático e intersección) para las clasificaciones binaria y ternaria y el conjunto de imágenes Insert (conjunto completo).	37
9. Resultados de la clasificación en % de ZMEG, aZIBO, B-ORCHIZ y ShapeFeat usando SVM con kernel intersección para los conjuntos de imágenes Insert-C, Insert-I (de izquierda a derecha) y para dos y tres niveles de desgaste.	46
10. Resultados de precisión en % de la combinación de ShapeFeat con B-ORCHIZ y otros descriptores como BCF (Bag of contour fragments), Hog (Histogram of oriented gradients) y SC (Shape Context)	48

Índice general

Agradecimientos	xv
1. Introducción	1
1.1. Motivación	1
1.2. Objetivos	2
1.3. Contribuciones principales	3
1.4. Organización de la tesis	4
2. Estado de la técnica	7
2.1. Monitorización del desgaste de la herramienta	7
2.1.1. Métodos indirectos	7
2.1.2. Métodos directos	11
2.2. Descriptores basados en forma	11
2.2.1. Parámetros básicos de forma	12
2.2.2. Métodos basados en funciones unidimensionales	13
2.2.3. Aproximación poligonal	15
2.2.4. Momentos	16
2.2.5. Dominios de transformada de forma	18
2.3. Descriptores basados en textura	22
2.3.1. Matrices de coocurrencia del nivel de gris	22
3. Conjuntos de imágenes	25
3.1. Conjunto de imágenes de recuperación de forma	25
3.2. Conjunto de imágenes del filo de corte de plaquitas	26
3.2.1. Recorte de los fillos de corte	28
3.3. Conjunto de imágenes de regiones de plaquitas	30
3.4. Conjunto de imágenes de plaquitas a alta resolución para descripción por textura	31
4. Descriptores basados en contorno	35
4.1. Metodología	35
4.1.1. CPDH	35
4.1.2. CPDH36R	36

4.1.3.	RCPDH	38
4.2.	Reconocimiento de forma	39
4.2.1.	Configuración del experimento	39
4.2.2.	Resultados	41
4.3.	Conclusiones	42
5.	Descriptores basados en momentos y orientaciones	45
5.1.	Metodología	46
5.1.1.	ZMEG	46
5.1.2.	aZIBO	47
5.1.3.	B-ORCHIZ	48
5.2.	Reconocimiento de forma	50
5.2.1.	Configuración del experimento	50
5.2.2.	Resultados	50
5.3.	Monitorización del desgaste de herramientas	52
5.3.1.	Configuración del experimento	52
5.3.2.	Resultados para el conjunto de imágenes de los bordes de plaquitas de corte	52
5.3.3.	Resultados para el conjunto de imágenes de regiones de plaquitas de corte	53
5.4.	Conclusiones	62
6.	Fusión de descriptores de regiones y momentos	65
6.1.	Metodología	65
6.1.1.	Descriptor de forma: ShapeFeat	65
6.2.	Monitorización del desgaste de herramientas	70
6.2.1.	Configuración del experimento	70
6.2.2.	Resultados para el conjunto de imágenes de regiones de plaquitas de corte	71
6.3.	Conclusiones	74
7.	Descriptor basado en parches	75
7.1.	Metodología	76
7.1.1.	Configuraciones de las regiones	78
7.1.2.	Descriptores de textura	79
7.2.	Monitorización del desgaste de herramientas	82
7.2.1.	Configuración del experimento	82
7.2.2.	Selección del umbral	83
7.2.3.	Configuración de la zona de desgaste y evaluación de la descripción.	84
7.3.	Conclusiones	86
8.	Conclusiones y perspectiva	87
8.1.	Resumen del trabajo	87

8.2. Contribuciones generales	88
8.3. Perspectiva	90
Lista de referencias	99

Anexo A: Actividades de investigación

Anexo B: Resumen de la tesis en castellano

Acknowledgements

A doctoral thesis is a difficult process and as time passes it becomes an arduous process, in which there are always times of downturn but also many good times, personal growth and motivation to go ahead and get the results that are sought. To carry it out, the support of people and institutions is always necessary.

Thanks to Junta de Castilla y León for letting me enjoy a grant destined to fund the hiring of recently graduated research personnel and the company TECOI for providing us an edge profile milling head tool and the inserts to create our inserts data set.

I am thankful to my supervisor Enrique Alegre from University of León for giving me the opportunity to carry out my doctoral studies under his supervision and guidance, for your patience and your advises and also to Joaquín Barreiro for your help.

I would like to express my special appreciation and thanks to Rocío Alaiz for co-supervising my thesis, encouraging my research, for cheering me up, for your collaboration and good advice and for making me see that I really want to dedicate myself to teaching.

Throughout these years I have enjoyed the company and support of good colleagues who have made this trip something incredible. Diego, thank you for your patience, for your help in making projects and for the good times you gave us. Thanks Laura, for all those conversations in the "Zielo"; Dani, for letting us be your lighthouse, Víctor and Claudia, for all the moments of laughter. Thanks also to María José and Fran for your support.

I would like to thank all my dear friends Eva, Patri, Janira, Jeny and also Andrea, Jessy, Jony, Aitor, Judit, Sabrina, Rubén, Rober, Nahúm and all the group for sharing all the good and dance and mountain moments, for your support, for your friendship and also for having patience when all the work has drift me apart from you.

Sobre todo, quiero dar las gracias a mis padres por haberme apoyado siempre en cada decisión, por la paciencia, por todo el cariño que me dais incondicionalmente y por cuidar de Limón y Pancho por mi. Ya se que os estáis esforzando mucho con

el inglés, pero quiero ponérselo fácil, que ya os toca.

Finally, my most heartfelt gratitude to you, Oscar. Had it not been for you, this would not have been possible. You're the only person who has given me all that I needed unconditional support in difficult times. You have become hard times in magical moments and this thesis is almost as much yours as mine. Thank you.

María Teresa García Ordás

León

12th July 2017

1.1. Motivation

In this dissertation, we address a real problem of tool wear monitoring using new techniques of computer vision based on shape and texture description. Some advantages of automatic systems are the avoidance of subjectivity against decisions or the improvement in terms of cost, therefore, the performance of the automated task.

Replacing tools at the right moment is essential because of the cost of cutting tools themselves, and also for the indirect costs due to the unproductive time needed to carry out the tool replacement. For this reason, tool condition monitoring has become a key point in industry for optimizing operations and increase productivity.

Developping solutions that can be implemented using low cost devices, quickly installable and with a very easy maintenance for automating processes, is also one of the most important goals of this work.

This research work has been conducted with the colaboration of TECOI, a company of Castilla y León specialized in the manufacture of advanced plate processing systems. The proposal will try to improve these machine tools by creating a method of detection of the milling cutting tool wear that facilitates the work of operators and also the reliability when making the cuts. One of the most important branches of production is the creation of structures of wind turbines, so this project also has a high impact on the priority area for the board of Castilla y León energy and the environment. Improving the production of these structures, they can create better wind turbines and thus multiply the amount of energy generated by them.

TECOI manufactures machine tools and one of its latest innovations has been the design and manufacture of a machine that solves a problem faced by metal poles of wind towers. Due to the process of cutting and weld preparation of the sheet, the staffs, manufactured by large curved plates and welded thick, present, after some time, micro-cracks. This problem has been eliminated with the TRF machine (see Figure 1.1) which cuts with plasma or oxyfuel but leaves welding profiled by milling the edge of the plate. This solution presents the problem of possible breakage of the cutting inserts at the end of the useful life. Evaluating their wear becomes crucial in

this field.

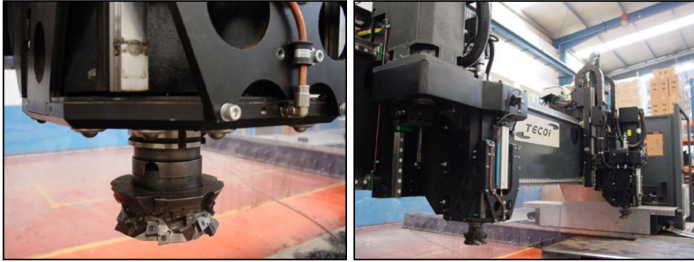


Figure 1.1: Left: overview of the machine. Right: detail of the head.

Determining the tool wear is a priority for the DPI2012-36166 project under which this thesis is framed. It will help prevent damage to the machine or workpiece by uncontrolled tool failure. The principal aim is to produce the final pieces in accordance with the requirements also proceeding to replace the cutting tool with a new one when it comes to its limit of usefulness. An example of a damaged insert can be seen in Figure 1.2.

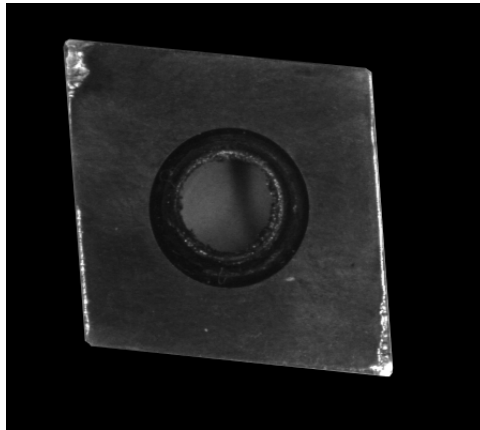


Figure 1.2: Example of a damaged insert.

1.2. Objectives

The main goal of this dissertation is to characterize the wear level of inserts in edge profile milling processes analyzing and clasifying digital images. Given the previous general goal, we defined the following particular objectives:

- To assess different digital image descriptors (shape, contour, moment and texture based) on wellknown and publicly available datasets.
- To propose new descriptors that cope with the weaknesses detected in the previous analysis.
- To create an insert dataset composed of multiples images which can reproduce the real scenario of tool wear monitoring.
- To categorize the tool wear using machine learning techniques and describing the wear region with the descriptors proposed in this work.
- To propose an online and automatic system that can be easily attached to all the different manufacturing tools in order to determine their insert wear.

1.3. Main contributions

The main contributions of this dissertation may be summarised as follows:

1. Three new datasets for tool wear monitoring have been created. **Insert edges dataset**, composed by 212 binary images, useful for contour and moment based descriptors which contains information about the shape of the wear, **Insert regions dataset**, composed by 573 images which distinguish the wear between the principal cutting edge of the insert and its adjacent edges. Finally, the **Insert High Resolution dataset** with 577 images developed for texture description analysis.
2. A new method based on CPDH, called CPDH36R has been proposed, implemented and tested using the classical datasets for shape retrieval: MPEG and Kimia. The proposal not only obtains better results in terms of hit rate but also outperforms the original CPDH in computational cost, reducing the processing time from days to just few seconds for large datasets.
3. A new descriptor called aZIBO has been proposed for shape description based on moments and rotational invariant features. This approach is based on the existing ZMEG algorithm but adds invariance to rotation which makes it more robust for all retrieving processes. The proposal has been also evaluated on the specific insert problem obtaining better results than the baseline ZMEG method.
4. A more complex evolution of aZIBO called B-ORCHIZ has been developed for insert wear evaluation. The feature description method adds new information connected with the boundary orientation which yields in a better description

approach not only for the insert task but also for the generic shape retrieval datasets.

5. An automatic clustering analysis based on dendrograms has been carried out to find the best way to divide the inserts by their wear level. The evaluation was done using B-ORCHIZ, aZIBO and ZMEG methods demonstrating the best discriminant power of the B-ORCHIZ algorithm in order to describe the inserts.
6. A study of the representative shape of wear classes has been performed in order to help the experts to easily identify the different levels comparing the new inserts with the representative shapes. A 2D and a 3D representation has been shown in order to clarify it better.
7. A new descriptor based on shape features of the inserts wear has been introduced. It is called ShapeFeat and it is a low computational cost descriptor that also requires a low quantity of memory to be stored. This proposal achieves great results in the insert wear classification using the Insert region dataset.
8. Fusion classification approaches have been used to take into account shape and contour information to classify the wear. We have demonstrated the good performance of ShapeFeat and B-ORCHIZ when they are merged.
9. A new online feature description method for automatic evaluation system of wear in milling processes has been proposed. This system avoids the segmentation step common to most description methods. It works on the basics of dividing the insert images into small patches isolatedly described. Classification method takes into account the proportion of damaged patches in the image to determine its class. It is a low cost approach, full automatic and easy to attach to all kind of cutting machines.

1.4. Thesis Organization

In this section the structure of this doctoral dissertation is described. The first introductory chapter has been focused on motivating the work presented in this dissertation, its main objectives and original contributions. The remaining chapters of the thesis are organised as follows.

A review of shape description and tool wear monitoring methods as well as a more specific revision of the state of the art for tool wear monitoring systems and specifically methods based on artificial vision systems is described in Chapter 2.

Chapter 3 presents the publicly available dataset on which our proposals were tested and also the new three datasets created in this dissertation: Insert

edges dataset, insert region dataset and insert high resolution dataset for texture description purposes.

Two new shape descriptors based on contour and moments are introduced in Chapter 4. These methods were tested on shape datasets demonstrating their effectiveness.

Moving on to the Chapter 5, two new description methods are explained: aZIBO and B-ORCHIZ that retain the advantages of both global and local descriptors.

Chapter 6 presents ShapeFeat method with the same goal of aZIBO and BORCHIZ: solving the problem of tool wear monitoring in edge profile milling processes. Using some existing methods, our proposal ShapeFeat was combined with the previous B-ORCHIZ obtaining still better results than using each method isolately.

The new feature description method based on dividing the image into patches is detailed in Chapter 7.

Chapter 8 contains a summary with the conclusions of this thesis and gives an outlook of possible future work lines to extend the presented work.

2.1. Tool Wear monitoring

In this dissertation, we address the task of tool wear description in edge profile milling processes using computer vision and machine learning techniques.

A tool condition monitoring system can be viewed as serving the following purposes: advanced fault detection system for cutting and machine tool, check and safeguard machining process stability, means by which machining tolerance is maintained on the workpiece to acceptable limits by providing a compensatory mechanism for tool wear offsets, and machine tool damage avoidance system (Dimla and Dimla, 2000).

It should be noticed that not only the cost of cutting tools is important, but also, and maybe much more important, indirect costs derived from the unproductive time needed to perform the tool replacement.

Optimizing tool replacement operations may produce a significant improvement in manufacturing efficiency and competitiveness. Conservative approaches usually replace the tools before it is actually necessary in order to avoid the high cost of tool failure, so they still have some lifetime. In (Malekian et al., 2009), it is estimated that cutting tool failures typically represent a 20% of the time a machine is inactive. However, replacing tools at the right moment is essential, not only because of the cost of cutting tools themselves, but also for the indirect costs due to the unproductive time needed to carry out the tool replacement. For this reason, tool condition monitoring has become a key point in industry for optimizing operations and increase productivity.

Approaches to tool wear monitoring proposed so far fall into two broad categories: indirect and direct methods. Next, a review of both of them is provided.

2.1.1. Indirect methods

Indirect monitoring methods estimate the wear by measuring variables such as tool temperature (Wang, Ming and Chen, 2016), cutting forces (Azmi, 2015; Wang, Yang, Xie and Zhang, 2014; Kaya et al., 2011), vibration (Scheffer and Heyns, 2001; Rao et al., 2014) or acoustic emission (Li, 2002), and also miscellaneous methods such

as ultrasonic and optical measurements, workpiece surface finish quality, workpiece dimensions, stress/strain analysis and spindle motor current, that are somehow correlated to tool wear stages. Next, a brief description of the proposals based on these techniques is provided.

Acoustic emission

Within the deformation zones, strain energy is released as the bonds between the metal atoms are disturbed. This released energy is commonly referred to as acoustic emission. Other sources of acoustic emission (AE) include phase transformations, friction mechanisms and crack formation or extension fracture. Acoustic sensors are cheap and nonintrusive, coupled with fast dynamic responses (Zhou et al., 2011) and for that reason, they have been used over the years to tool wear monitoring. For example, in (Zhou et al., 2011), the authors use the acoustic emission from an embedded sensor for computation of features and prediction of tool wear. The metal cutting process usually initiates with the occurrence of plastic deformation of workmaterial and is followed by tear and removal of material from the workpiece. This process ultimately damages cutting tool and causes tool wear. In (Bhuiyan et al., 2016), an acoustic emission (AE) sensor has been employed to measure the signal frequency in machining. The frequency of tool wear for this particular investigation is found to lie between 67 kHz and 471 kHz whereas for plastic deformation of workmaterial, it has a fluctuation within the range starting from 51 kHz to some value within 471 kHz. In the same way, in (da Silva et al., 2016), a new parameter based on acoustic emission signal energy (frequency range between 100 and 300 kHz) was introduced to improve response. Tool wear in end milling was measured in each step, based on cutting power and acoustic emission signals. The wear conditions were then classified taking into account two classes, the signal parameters were extracted, and then, a probabilistic neural network was applied as classifier.

Tool cutting edges temperature

Metal cutting generates a significant amount of heat. The resultant high temperatures around the cutting tool edges affects directly on the rate and mode of cutting tool wear, the friction between chip and cutting tool, and also that between the cutting tool and the newly formed surface. Frictional behaviour on the tool faces is thought to affect the geometry of the cutting process. In (Ceau et al., 2010), an experimental study on temperature measurement in turning operation was carried out and more recently, in (Chryssolouris et al., 2012), a methodology is described to conduct temperature measurement on the cutting edges and the clearance faces on twist drills using a fibre optic two color pyrometer. The authors

used two measuring positions of the fibre in order to determine the temperature at two different locations, centre and outer corner of the drill. The measurements were carried out on a stationary work piece and a rotating drill.

Unfortunately, the existing tool wear models are all established to predict tool wear in continuous turning process, not suitable for the interrupted milling process. To solve this problem, in (Wang, Ming and Chen, 2016), the authors firstly determines the threshold temperature of adhesive wear, and then, a mathematical model of milling tool's flank wear is then proposed by acquisition of the real time cutting force and temperature.

Cutting forces

It has been widely established that variation in the cutting force can be correlated to tool wear (Wang, Yang, Xie and Zhang, 2014; Azmi, 2015; Kaya et al., 2011). Mehdi et al. (Nouri et al., 2015), developed a new method to monitor end milling tool wear in real-time by tracking force model coefficients during the cutting process. The behavior of these coefficients are shown to be independent from the cutting conditions and correlated with the wear state of the cutting tool. Experiments have shown that the three components of the cutting force (Figure 2.1) respond differently to the various wear forms occurring on the tool. For example, the feed (F_x or F_f) force is insensitive to crater wear whereas the feed and radial (F_y or F_r) forces may be influenced more by tool wear than the main cutting (F_z or F_t)

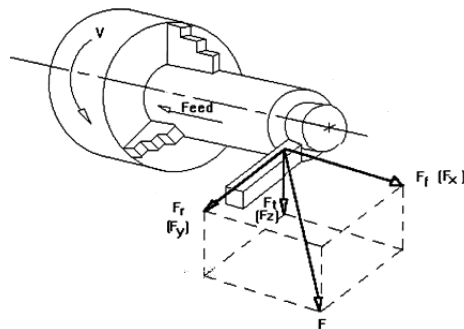


Figure 2.1: Cutting force components on a single point during the turning process.

Vibration signatures

Vibrations are produced by cyclic variations in the dynamic components of the cutting forces. Usually, these vibrational motions start as small chatter responsible

for the serrations on the finished surface and chip thickness irregularities, and progress to what has come to be commonly termed vibration. Mechanical vibrations generally result from periodic wave motions. Direct measurement of vibration is difficult to achieve because its determining characteristic feature, the vibration mode is frequency dependent.

In (Bovic Kilundu, 2011), the authors use signals in order to determine the tool wear. Pseudo-local Singular Spectrum Analysis (SSA) is performed on vibration signals measured on the tool holder. This is coupled to a band-pass filter to allow definition and extraction of features which are sensitive to tool wear. These features are defined, in some frequency bands, from sums of Fourier coefficients of reconstructed and residual signals obtained by SSA.

Also in (Prasad and Babu, 2017), a correlation between vibration amplitude and tool wear when in dry turning is analyzed via experiments and finite element simulations. In this work, the primary concern is to find the relative vibration and tool wear with the variation of process parameters. These changes lead to accelerated tool wear and even breakage.

Fused signals

An extensive and elaborate experimental investigation into the development of an on-line tool wear monitoring system for metal turning operations using cutting force measurements fused with vibration signatures was developed in (Dimla, 1999).

Also, a multiple sensor monitoring system comprising cutting force, acoustic emission and vibration sensing units was employed in (Segreto et al., 2013) for tool state assessment during turning of Inconel 718 nickel alloy. Feature extraction was realised by processing the detected sensor signals in order to reduce the high dimensionality of the sensorial data. To improve manufacturing system reliability, in (Wang et al., 2017), a new virtual tool wear sensing technique based on multisensory data fusion and artificial intelligence model for tool condition monitoring is presented. It infers the difficult-to-measure tool wear parameters (e.g. tool wear width) by fusing in-process multisensory data (e.g. force, vibration, etc.) with dimension reduction technique and support vector regression model.

But these indirect methods based on signals and sensors do not provide the sufficient information to perform an optimal inserts replacement because the relationship between tool wear and the observed variables depends on the cutting conditions and, in general, it is not known in advance. Although these methods are the most popular, the precision achieved with them is seriously affected by noise signals in industrial environments. Also, they are cost-inefficient and unavailable for real applications (Kassim et al., 2007).

2.1.2. Direct methods

In contrast to indirect methods, direct methods - usually based on computer vision systems - have the advantage of measuring actual geometric changes in the tool, offering more accuracy and reliability (Loizou et al., 2015). A review of machine vision sensors for tool condition monitoring is presented in (Kurada and Bradley, 1997; Dutta et al., 2013).

Some of these methods are based in the wear shape contour, others are based on the texture of the wear, shape properties or combination of some of them. In the following two sections, a brief review of the state of the art of shape and texture methods is summarized in order to put in context the contributions of these dissertation.

2.2. Shape based descriptors

In the last decades, the fast increase of image capture devices has made the number of digital images to grow exponentially. Therefore there is a growing demand of techniques for indexing, storing and retrieving these digital images.

In this section we show some examples that can be adapted to our particular problem of tool wear monitoring.

Many of the methods describing shapes have been developed over the years. Images are usually described by features based on contour, color, shape or texture. All of the following methods of shape description can be applied on our particular problem: tool wear monitoring.

Two of the key factors in an image retrieval system are its accuracy and its computational efficiency. Image retrieval using shape information has received some attention by the image processing and pattern recognition community, focusing on improving any, or even both of these factors. Some works focus on carrying out the recognition by using distance measures between the shapes. For example, Nasreddine et al. defined a multi-scale distance based on geodesics in the shape space which is robust to outliers and invariant to translation, rotation and scaling (Nasreddine et al., 2010). Ćurić et al. (Curic et al., 2014) proposed other distance measure called complement weighted sum of minimal distances, which can be used for shape registration and matching. In their proposal the contribution of each point of each set is weighted according to its distance to the complement of the set.

There are several groups of methods for shape representation depending on the information extracted to build the descriptor. In the survey of Mingqiang et al (Mingqiang et al., 2008) seven different types of shape representation are presented: Basic shape parameters methods, one-dimensional function methods, polygonal

approximation, spatial interrelation features, moments, scale-space methods and finally shape transform domains.

2.2.1. Shape parameters

Shape representation based on basic shape parameters usually can only discriminate shapes with large differences. However, there are multiples fields where using this kind of features is very useful as filters to eliminate false hits or combined with other kind of features. Circularity ratio represents how a shape is similar to a circle. In (Wang, Chang, Zhou, Li, Liu and Guo, 2016) this ratio is used to identify the five different types of white blood cells evaluating their shape in images captured with a microscope. This ratio can be defined as the ratio of the area of a shape to the area of a circle having the same perimeter. Another basic shape feature widely used in the literature is the convexity. It is defined as the ratio of the perimeter of the convex hull P_{Ch} over that of the original contour P (see equation (2.1)).

$$Convexity = \frac{P_{Ch}}{P} \quad (2.1)$$

We can define a convex region if for any two points $P_1, P_2 \in R^2$ the whole line segment P_1P_2 is inside the region. Taking that into account, the convex hull of a shape is the smallest convex region including it. Figure 2.2 shows an example.



Figure 2.2: Illustration of convex hull

In the work carried out by Yasseen et al. (Yasseen et al., 2016) they proposed a method for shape retrieval based on dividing the image into visual parts and describing them using shape parameters such as their area, circularity, convexity or solidity.

Recently, Fishbaugh et al. (Fishbaugh et al., 2017), proposed a model which decouples the deformation parameters from the specific shape representations, allowing the complexity of the model to reflect the nature of the shape changes, rather than the sampling of the data.

2.2.2. One dimensional function methods

One dimensional function methods usually capture the perceptual feature of the shape. Although these features can be used to describe a shape just by themselves, it is frequently used in combination with other feature extraction methods like fourier or wavelet descriptors.

In (Condorovici et al., 2015), the authors extensively mimic low-level and medium-level human perception by relying on perceptually inspired features. In their work, while Gabor filter energy has been used for art description, Dominant Color Volume (DCV) and frameworks extracted using anchoring theory are novel in this field. For the classification, a combination of Multi-Layer Perceptron (MLP) with Support Vector Machines (SVM) was performed.

Curvature is also a very important feature for human to discriminate between shapes. It also has salient perceptual information and has proven to be very useful for shape recognition. In (Czerniawski et al., 2016), the authors present an automated method for locating and extracting pipe spools in cluttered point cloud scans. The method is based on local data level curvature estimation, clustering, and bag-of-features matching. Figure 2.3 is an example of curvature function.

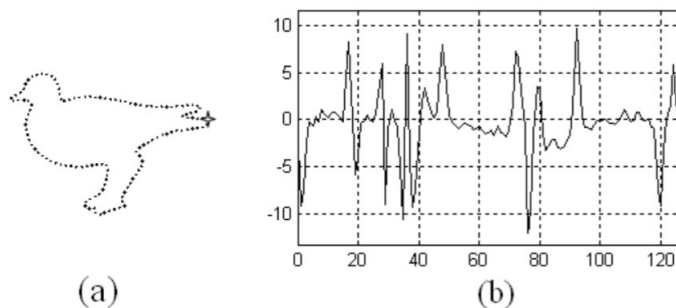


Figure 2.3: (a) Contours normalized to 128 points. The star indicate the starting point. (b) curvature function computed in clockwise.

Most of the reported descriptors are still facing accuracy and computational challenges. Fourier descriptors are considered to be promising descriptors as they are based on a sound theoretical foundation and also have the advantages of computational efficiency and attractive invariance properties. Based on this information, Elghazal et al. (El-ghazal et al., 2012), proposed a new curvature-based Fourier descriptor (CBFD) for shape retrieval. The proposed descriptor takes an unconventional view of the curvature-scale-space representation of a shape contour as it treats it as a 2D binary image (hence referred to as curvature-scale image, or CSI). The invariant descriptor is derived from the 2D Fourier transform of the curvature-scale image. This method allows the descriptor to capture the detailed dynamics of the shape curvature and enhance the efficiency of the shape-matching process.

Another interesting method widely used in shape representation is the area function. When the boundary points change along the shape boundary, the area of the triangle formed by two consecutive boundary points and the centroid also changes. With this, an area function can be formed in order to exploit it as shape representation (see Figure 2.4). A variant of this area function is the Triangle Area Representation (TAR) which is computed from the area of the triangles formed by points in the shape boundary. TAR is invariant to translation, rotation, scaling and robust against noise and moderate occlusion.

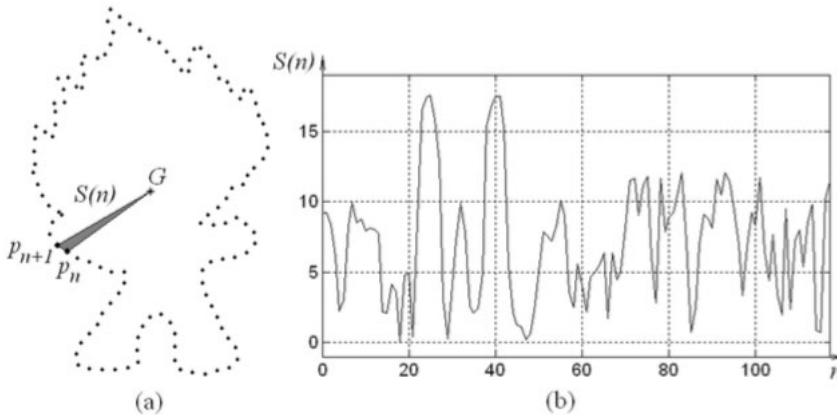


Figure 2.4: (a) Original contour. (b) the area function of (a).

All of these methods based on one-dimensional function derived from shape contour. To obtain the translation invariant property, they are usually defined by relative values. The invariance to scale is easily obtained normalizing the final descriptor. Orientation changes are usually avoided in the matching process by

shifting the descriptors until reach the shortest distance. However, they are sensitive to noise and slight changes in the boundary points can cause large errors in the overall description. Usually these kind of descriptors are combined with further processing to increase it robustness and reduce the matching load.

2.2.3. Polygonal approximation

Polygonal approximation methods are used to decrease the variation along the edge in the shape, trying to capture the overall shape information. These techniques are very useful because they reduce the effects of discrete pixelization of the contour. In general, there are two different methods to do it: Merging and splitting.

Merging methods add successive pixels to a line segment if each new pixel which is added does not cause the segment to deviate too much from a straight line. The distance threshold method starts by choosing two points from the boundary and calculating the squared error for every point of the boundary between those two selected points. If the error exceeds a threshold, we choose two nearer points to create the segment and start again the process. Wang et al in (Wang, Feng, Bai, Liu

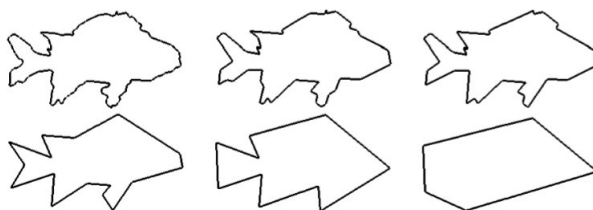


Figure 2.5: A few stages of polygon evolution according to a relevant measure.

and Latecki, 2014) proposed a method based on a bag of contour frames extracted from the original shapes using this polygon evolution method.

On the other hand, splitting methods work by first drawing a line from a point on the boundary to another. After that, we compute the perpendicular distance from each point along the boundary and if this exceeds some threshold, we break the line in two segments at the point of greatest distance, repeating the process until we do not need to break any more (see Figure 2.6). In (Madrid-Cuevas et al., 2016), both strategies are combined (splitting and merging). The proposed algorithm obtains polygonal approximations of contours and does not have parameters that must be adjusted, which provides the best balance between fidelity and efficiency, and has a modest algorithmic complexity. The method is based on an analysis of the convexity and concavity tree of the contour, and an efficient split/merge strategy is used. In the last years, some methods based on salience points (Kim and Park, 2014)

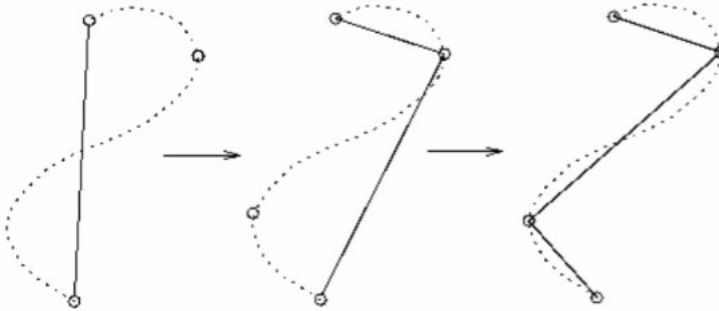


Figure 2.6: Splitting method for polygonal approximation.

and shape curvature for describing an image have been developed. In (Andaló et al., 2010) IFT is used but the authors also introduce the concept of tensor scale, which is a morphometric parameter that unifies the representation of local structure thickness, orientation and anisotropy. (IFT is a generalization of Dijkstra algorithm: it is a graph based approach which returns a root map, a cost map and a label map containing the relevant information of the contour points and its relationship with its influence area points.). Also salience points techniques is used in (Pedrosa et al., 2013). The technique presented here consists of: a salience point detector; a salience representation using angular relative position and curvature value analyzed from a multi-scale perspective; and a matching algorithm considering local and global features to calculate the dissimilarity. The technique proposed in (Pedrosa et al., 2013) is robust to noise and presents good performance when dealing with shapes of different classes but visually similar.

Recently, Laiche et al. (Laiche et al., 2014), proposed a novel part-based approach for two dimensional (2-D) shape description and recognition. According to this method, first the polygonal approximation is employed to represent the outline shape by an ordered sequence of parts. Then using the Least squares model, each part is associated with a cubic polynomial curve. The obtained curves are normalized that are invariant to scaling, rotation and translation. Finally, based on shape similarity of resulting curves, a shape similarity between an input shape and its reference model is defined. The authors also proposed a two-step matching algorithm.

2.2.4. Moments

The concept of moment in mathematics evolved from the concept of moment in physics. For contour and shape description, the moment theory can be used to

analyze the objects. We can define the r^{th} moment m_r and central moment μ_r of a 1-D shape representation $z(i)$ as it is shown in equations (2.2) and (2.3)

$$m_r = \frac{1}{N} \sum_{i=1}^N [z(i)]^r \quad (2.2)$$

$$\mu_r = \frac{1}{N} \sum_{i=1}^N [z(i) - m_1]^r \quad (2.3)$$

where N is the number of points in the boundary. The advantage of this description techniques is that it is very easy to implement but sometimes it is difficult to associate higher order moments with physical interpretation.

Furthermore, region moments are becoming very popular since the mid-20th century. In 1962, Hu proposed invariant moments to scale, rotation and translation (Hu, 1962). However, this descriptor had a lot of redundancy and it was very noise sensitive.

The algebraic moment invariants are computed from the first m central moments and are given as the eigenvalues of predefined matrices, $M_{[j,k]}$, whose elements are scaled factors of the central moments.

Zernike moments (Khotanzad and Hong, 1990) are orthogonal moments. The complex zernike moments are derived from orthogonal Zernike polynomials. These moments have the following advantages:

- Rotation invariance
- Robustness to noise and minor variations of the shape
- Low redundancy

However, the extraction of Zernike moments pose several drawbacks:

- The image space must be transformed to the domain where the orthogonal polynomial is defined.
- The continuous integrals must be approximated by discrete summations.
- The computational complexity increases as the order becomes large.

These moments are still widely used today.

In 2016, a very interesting work (Wu et al., 2016) was carried out to detect students action through a camera. The method is based on Zernike moment description of the image to identify three different student status: Hands up, Stand up and Sit Down. Furthermore, this technique has been combined with a Lucas-Kanade optical flow to describe the direction and speed of the movement.

Results demonstrated the effectiveness and robustness of the method identifying all the three different actions.

At the beginning of 2017, Kumar et al. proposed a method to identify plants based on leaf images. In order to deal with this challenging problem, they combined shape and texture descriptors. An orthogonal moment based descriptor called Krawtchouk moments was used as shape descriptor whereas HOG and Gabor features were implemented as texture description. The system was trained by using a Support Vector Machine learning algorithm achieving high classification rates (Pradeep Kumar et al., 2017).

Recently, (Bolourchi et al., 2017) describe standard synthetic aperture radar (SAR) images using Radial Chebyshev moments fused with Zernike moments. Their experimental results demonstrate that RCM improves Zernike moments description in a 10% of hit rate. However, the fusion of both moment descriptors increases the total accuracy of the classification in more than a 6%.

All of these studies show that moment based descriptors are usually concise, robust and easy to compute. Moreover, most of them are invariant to rotation, scaling and translation. However, because of their global nature, these methods have problems trying to correlate high order moments with salient features of the shapes.

2.2.5. Shape transform domains

Finally, shape transform domains include methods which are formed by transforming the detected objects or the whole image. These transformations can be used to describe and characterize the appearance of the shapes being represented by all or partial of the transform coefficients. Fourier descriptor was developed more than 40 years ago but nowadays it is still considered a valid description tool. It is easy to compute, robust to noise and very compact. Usually, Fourier descriptor is obtained by applying Fourier transform on a shape signature that is a one-dimensional function which is derived from a shape boundary coordinates. The most efficient one is the Fourier descriptor which derives from the centroid distance function $r(t)$ given by equation(2.4)

$$a_n = \frac{1}{N} \sum_{t=0}^{N-1} r(t) \exp\left(\frac{-j2\pi nt}{N}\right), \quad n = 0, 1, \dots, N - 1 \quad (2.4)$$

In (Sokic and Konjicija, 2016), the authors propose a novel method of extracting Fourier descriptors that preserve the phase of Fourier coefficients. They introduce specific points, called pseudomirror points, and use them as a shape orientation reference. These points facilitate the extraction of phase-preserving Fourier descriptors which are invariant under translation, scaling, rotation and starting

point change. In this same line, we can find more and more recent works still using Fourier with satisfactory results ((Shu et al., 2015; Ai et al., 2013)).

Another important transform method is the Wavelet transform. This descriptor decomposes a curve into components of different scales so that the coarsest scale components carry the global approximation information while the finer contains information about the details. This technique has a lot of interesting properties such as invariance to scaling, translation and rotation, multi-resolution representation, stability, uniqueness and spatial correlation. In (Hadizadeh, 2015), a given gray-scale texture image is first processed by a bank of Gabor wavelets (filters) at different scales and orientations. The obtained filters' responses are then further processed, and a set of local binary patterns called "Local Gabor Wavelets Binary Patterns" (LGWBPs) are computed by comparing the local filters' outputs at different orientations with the global mean of filters' outputs at the same orientations. The obtained patterns are then converted to a number of decimal rotation-invariant codes, and a histogram of the resultant codes at different scales is finally used as a texture feature vector. Although Wavelet is an old method, it is still in use just like Fourier ((Shahsavari et al., 2017; Masoumi et al., 2016))

Some researching groups have merged Fourier and Wavelet transforms in order to develop a more complex and efficient method. For example, in (Yuan et al., 2014), Fourier and Wavelet transformation were adopted to analyze shape characteristics. The Fourier descriptors can be used to characterize the shape from the aspects of the first eight Normalization amplitudes, the number of the largest amplitudes to inverse reconstruction, similarity of shapes and profile roughness. And the Diepenbroek Roughness was rewritten by Normalization amplitudes of FFT (Fast Fourier Transform) results. Moreover, Sum Square of Relative Errors (SSRE) of Wavelet transformation (WT) signal sequence, including approximation signals and detail signals, was introduced to evaluate the similarity and relative orientation among shapes. As a complement to FFT results, the WT results can retain more detailed information of shapes including their orientations. Besides, the geometric signatures of the second phases were extracted by image processing and then were analyzed by means of FFT and WT.

Other more recent works deal with computing shape or boundary descriptors to carry out the recognition by means of similarity measures or other classification methods. Biswas et al. presented a shape indexing and retrieval method which scales up for large databases, while it performs as well as other state-of-the-art methods (Biswas et al., 2010). This method is invariant to articulations (which made it useful for human pose estimation tasks) rigid transformations, etc., and is formed by geometric relations between pairs of interest points of the shape. Shu and Wu et al. (Shu and Wu, 2011a) introduced a shape contour descriptor for shape matching and retrieval purposes, which was based on the distribution of points on the object contour in polar coordinates. Its computational complexity is low (which

makes it feasible for practical applications) and it is invariant to scale. However, it can only deal with images with a single closed contour, and it is not invariant to rotation. It is invariant to scale and translation and the problem of the invariant to rotation can be partially resolved in the matching process. After the description of images is generated, the similarity value of the images is obtained by EMD (Earth Mover's Distance) metric. In order to make the EMD method used effectively for the matching, they also developed a new approach to the ground distance used in the EMD metric under polar coordinates. Experimental results of image retrieval demonstrate that their proposed descriptor has a strong capability in handling a variety of shapes. This proposal is one of those that we base on to try to obtain better results.

Another shape based retrieval method that is indeed rotation and scale invariant is presented in (Mukanova et al., 2014), based on Gestalt principles of human vision. In (Zagoris et al., 2011), Zagoris et al. proposed a descriptor containing conventional contour and region shape features with a very low size and a low computational cost which achieved good retrieval accuracy in databases of handwritten words and the MPEG-7 dataset. Hu et al. (Hu et al., 2012) proposed two Perceptually Motivated Morphological Strategies (PMMS) to enhance the retrieval performance of common shape matching methods, based on carrying out morphological operations on the shapes. Afterwards, the same authors proposed two new shape descriptors, called Angular Pattern (AP) and Binary Angular Pattern (BAP), and a multi-scale integration of both for shape retrieval (Hu et al., 2014). Both are global shape descriptors and invariant to scale and rotation. Laiche et al. (Laiche et al., 2014) proposed a descriptor of 2D objects based on the analysis of their boundaries. It is based on decomposing the shape of the object in a sequence of parts using a polygonal approximation, and then approximating each part by means of a polynomial curve. Finally, the representation is transformed into invariant features. The results showed the invariance of the approach to some affine transformations.

However, classic methods as Zernike moments are still being used as shape descriptor with good results. For example, Anuar et al. (Anuar et al., 2013), use Zernike moments as global descriptor combined with a local descriptor called EGCM and Wei et al. (Wei et al., 2009) extract the curvature and centroid distance after the edge extraction as well as image Zernike moments.

Although Zernike moments are widely used, they can only be applied on gray scale images, so methods for color images were developed with the same proposal of image retrieval. In (Zhu et al., 2013), the authors propose a new operator called the Orthogonal Combination of Local Binary Patterns (OC-LBP) and six new local descriptor based on OC-LBP enhanced with color information for image region description. Proen and Santos also proposed a method for recognition of degraded iris images acquired at visible wavelengths (Proen and Santos, 2012).

Nowadays the combination of global and local descriptors is widely used. One

example is the work developed by Singh et al. (Singh and Pooja, 2012), who used the angular radial transform (ART) as global descriptor and polar Hough transform (PHT) as local one. In (Anuar et al., 2013), a combination of global and local descriptors is also used but in this case, they employ Zernike and Edge Gradient Co-occurrence Matrix (EGCM) as global and local descriptors, respectively.

Although all of these proposals of retrieval system can be used in tool wear monitoring, it can be also applied in other fields. For example, in (Blenkinsopp et al., 2012), a method for measuring dynamic dorsal foot surface shape and deformation during linear running using digital image correlation is proposed and Hao et al. (Hao et al., 2013) carried out an active learning based intervertebral disk classification combining shape and texture similarities.

2.2.6. Shape based description in tool wear monitoring

Contour methods have some limitations since do not provide all the necessary information to determine the wear level because there are many other factors which characterize it than only its shape. Recently, the computer vision research community has successfully developed new methods which use shape information in order to describe the images of the inserts. These shape features are extracted from the images and later on processed by a classification or regression model to provide relevant information of the tool wear status. In (Dominguez-Caballero et al., 2016), a new method to describe the wear of a ceramic cutting tool (SiAION) has been carried out by firstly segmenting the spark area with techniques based on low and high pass filters. After that, a description based on a connected component discrimination of the shape of the region was performed.

Shape description methods have already been applied in the tool wear monitoring field to ensure the optimal replacement of the tool in lathe processes and in this line, Barreiro et al. propose the use of Hu and Legendre moments (Barreiro et al., 2008).

Our problem is quite different from that of (Barreiro et al., 2008) in the sense that we determine the tool wear in edge profile milling processes instead of lathe processes. In milling, the wear shape is distributed uniformly along the insert while in lathe processes the wear occurs on the same area concavely. Also, most works that deal with milling focus on end milling or face milling. Thus, monitoring tool wear in edge profile milling processes states a more challenging task since the insert distribution is not typical and the head geometry is more complex.

2.3. Texture based descriptors

Nowadays, if we focus on digital image processing, texture methods are one of the most widely used techniques. Texture features are usually computed from the statistical distribution of the intensities of all the pixels with respect an specific position relative to each other in the image. In the following sections, we present different ways to extract those features and their applicability in tool wear monitoring processes.

2.3.1. Gray Level Coocurrence Matrices

The Gray Level Coocurrence Matrix (GLCM) method is a way of extracting very useful information of second statistical order. A GLCM is a matrix where the number of rows and columns is equal to the number of gray levels in the image (usually 256). The element of the matrix $M(i,j)$ is the relative frequency with which two pixels, separated by a fixed distance and orientation, occur within a neighborhood, one with intensity i and the other with intensity j .

Plenty of different works have used this way of extracting information to build their feature descriptors.

In 1973, Haralick and his team developed a method (Haralick et al., 1973) based on features extracted over the GLCM and evaluated it over three diferent datasets of sandstones, aerial photographs and Earth Resource Technology Satellite (ERTS) multispecial imagery. In their work, Haralick et al. extracted 14 different features for the GLCM of an image: Angular speed moment, Contrast, Correlation, Sum of Squares, Inverse Difference Moment, Sum average, Sum variance, Sum entropy, Entropy, Difference Variance, Difference Entropy, Information Measures of Correlation and Maximal correlation coefficient. Results demonstrate a high performance and nowadays, the Haralick features extracted from the GLCM are still widely used for texture description. In (Zayed and Elnemr, 2015), Zayed and Elnemr proposed a new evaluation system based on these Haralick features to discriminate lung abnormalities.

In 1980, Connors and Harlow demonstrated that GLCM is a more powerfull technique than Gray Level Difference Matrix (GLDM), Gray Level Run Length Method (GLRLM) and the Power Spectral Method (PSM) (Connors and Harlow, 1980)

More recently, Mohanaiah et al. (Mohanaiah et al., 2013) proposed a new approach of GLCM feature extraction based on the evaluation of just four features, which improves not only the processing speed but also the accuracy on classification of cartoon images.

2.3.2. Local Binary Patterns

Local Binary Patterns (LBP) was proposed by Ojala et al. in (Ojala and Pietikainen, 1996). In this work, Ojala and Pitikainen present a new method based on the extraction of a binary code for each pixel taking into account its neighborhood gray level intensity, assigning 1 if the neighbor intensity level is higher than the central pixel and 0, otherwise. After that, a histogram is created to represent the frequency of all the different binary code possibilities which forms the LBP feature descriptor. In (Ojala and Pietikainen, 2002), the same authors proposed two extensions of the original operator. First of all, they defined LBP for different neighborhoods configuration and secondly, they defined the so-called uniform patterns, an LBP binary code which less than three value changes. For example, 00001111 or 11000001 are uniform patterns whereas 10010100 is not.

LBP has been widely used in the literature due to its low computational cost and its high description power. In (Doost and Amirani, 2013), a new method which combines LBP and Continues Wavelet transformation has proposed for texture classification over a textile dataset with 12 different classes divided to 64 images. In (Ahonen et al., 2006), Timo Ahonen and his team, developed a method for face description using LBP achieving better results than with other texture description methods such as Difference of histograms or Texton Histogram, in a lower processing time. In the work carried out by Garcia-Olalla et al. (García-Olalla et al., 2015), a new method for determining the integrity of the acrosome in boar spermatozoa has been carried out by combining LBP with other texture and shape descriptors achieving almost a 100% of success rate.

In the methodology section we are going to go deeper with LBP and some of their most important variations.

2.3.3. Texture based description in tool wear monitoring

In (Datta et al., 2013), Datta et al. proposed a method based on texture analysis and Voronoi tessellation to measure tool wear.

In (Dutta et al., 2012), online acquisition of machined surface images has been done. Then, these images were analysed using an improvised grey level co-occurrence matrix (GLCM) technique with appropriate pixel pair spacing (pps) or offset parameter. After this work, the authors proposed to employ the discrete wavelet transform on turned surface images (Dutta et al., 2016b) and also texture analysis and support vector regression (Dutta et al., 2016a).

In (Li and An, 2016), a new method also based on Gray Level Co-occurrence Matrix (GLCM) has been proposed. They use it to extract the best texture analysis window of an insert, showing that entropy remains monotonic as tool wear increases, demonstrating that entropy can be used effectively to monitor

tool wear. GLCM is also used in (Sun et al., 2017) where the authors propose an in-process surface texture condition monitoring approach. Based on the grey-level co-occurrence matrices, some surface texture image features are extracted to describe the texture characteristics.

In (García-Olalla et al., 2017), several descriptors have been extracted based on the LBP from the Outex dataset, as well as their combination with LOSIB (Local Oriented Statistical Information Booster) but in this case, all descriptors achieved a lower performance level combined with LOSIB because Outex is a data set that studies rotation invariances.

Recently, in (Antić et al., 2018), the module of the Short Term Discrete Fourier Transform (STDFT) spectra is considered. It is obtained from the particular vibration sensors signal utterance as the 2D textured image. This is done by identifying the time scale of STDFT as the first dimension, and the frequency scale as the second dimension of the particular textured image. They do not use a photography system but take into account the power of texture description applied over the vibration forces.

In this chapter we present and describe the datasets that were used in this dissertation in order to evaluate our proposals.

3.1. Shape retrieval datasets

Three classic shape datasets were used to test some of the descriptors proposed in this work: Kimia25 and Kimia99 datasets, and a subset of the MPEG7 dataset composed of 400 images from MPEG7 called MPEG2. Kimia25 is composed of 6 classes. Each of them contains four shapes except the “hand class” which is composed of 5 samples. Kimia99 is composed of 9 classes, each one containing 11 shapes. Both Kimia25 and Kimia99 datasets are illustrated in Figure 3.1. MPEG7 contains 70 classes with 20 samples for each of them. MPEG2 is an MPEG7 subset composed for 20 of the classes of MPEG7 dataset. This MPEG2 subset was created due to the high computational cost of CPDH, which spends more than a week to carry out the classification experiments. Some examples of the shapes contained in MPEG2 dataset are shown in Figure 3.2.

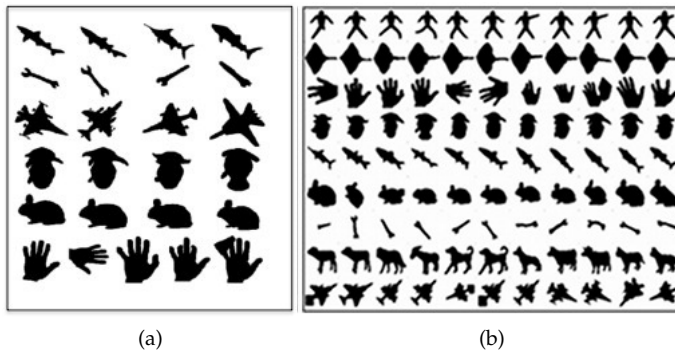


Figure 3.1: a) Kimia25 database composed by 25 samples divided in 6 classes. b) Kimia99 database composed by 9 classes with 11 samples per class.

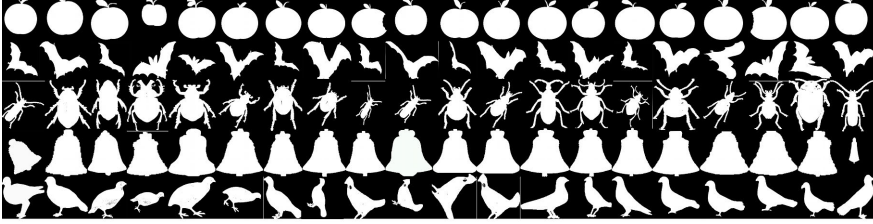


Figure 3.2: Example of 4 classes with 20 samples of MPEG2 dataset.

3.2. Insert edges dataset

One of the goals of this work is to characterize the insert state based on the wear region shape. For this purpose, we have acquired an Insert edges dataset, processing the images of 53 cutting tools as the ones shown in Figure 3.3.

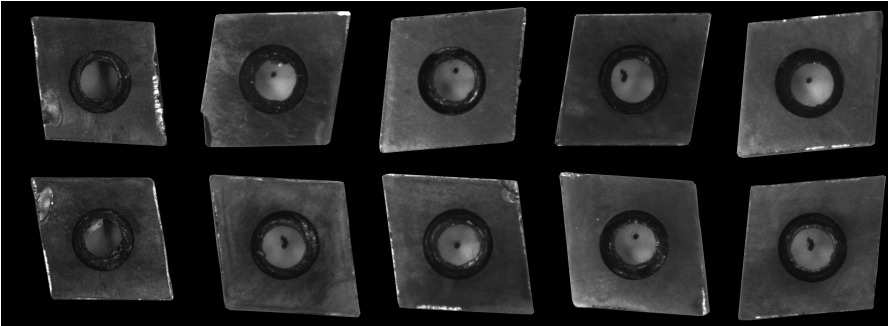


Figure 3.3: Insert dataset illustration. Inserts on the left present high wear while inserts on the right present low wear taking into account the left cutting edge.

Using grayscale images, instead of 3D images (as other authors do), we take care of measuring the wear of the insert. Following this approach we are able to monitor the wear in real time what makes feasible the deployment of our proposal in a real environment.

The images were captured using a monochrome camera, model Genie M1280 1/3" with a 25mm optic AZURE. The focus and aperture are manual. The sensor has a resolution of 1280×960 pixels. The monochrome camera and its specifications are shown in Figure 3.4 and in Table 3.1.

In order to improve the contrast in black and white image, two lighting bars of red LEDs (BDBL-R8216H) were used. Images have been taken on the inserts dismantled from the cutting head and placed on an uniform background.

To acquire the images, a prototype in which the camera was placed in a support

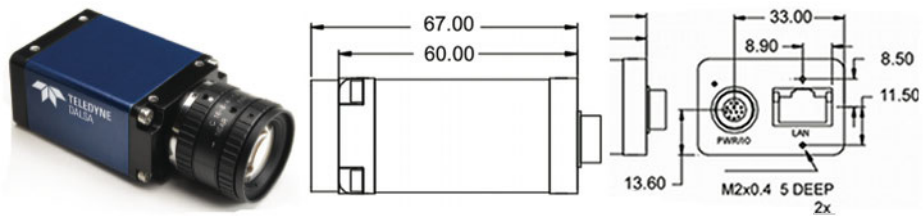


Figure 3.4: Genie M1280 (left side), lateral view (center image), back view (right side).

Table 3.1: Genie M1280 specifications.

Camera Specifications	
Active Resolution	1280 x 960
Frame Rate	24fps
Exposure Control	Programable, or via External Trigger
Output	Gigabit Ethernet, also support 100 Mbps (RJ-45)
Lens Mount	C or CS-mount / Right Angle available
Size	44mm x 29mm x 67mm (including lens adapter)
Power Supply	12 to 24 v (4 W)

with adjustable distance, was constructed. This prototype and also the lighting system are shown in Figure 3.5.

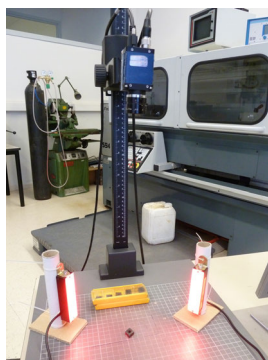


Figure 3.5: Image capture prototype. The support in which the camera is placed and the LEDs bars employed are shown.

3.2.1. Cropping of the cutting edges

Each insert, captured as a grayscale image and with a masked background (Figure 3.6 (a)) is subjected to a pre-processing step that yields four new images, one for each of the four cutting edges of the insert. Therefore, we crop each cutting edge and rotate the two vertical ones, obtaining four images with a horizontal cutting edge on it, for each insert processed. In this section, this pre-processing step is explained.

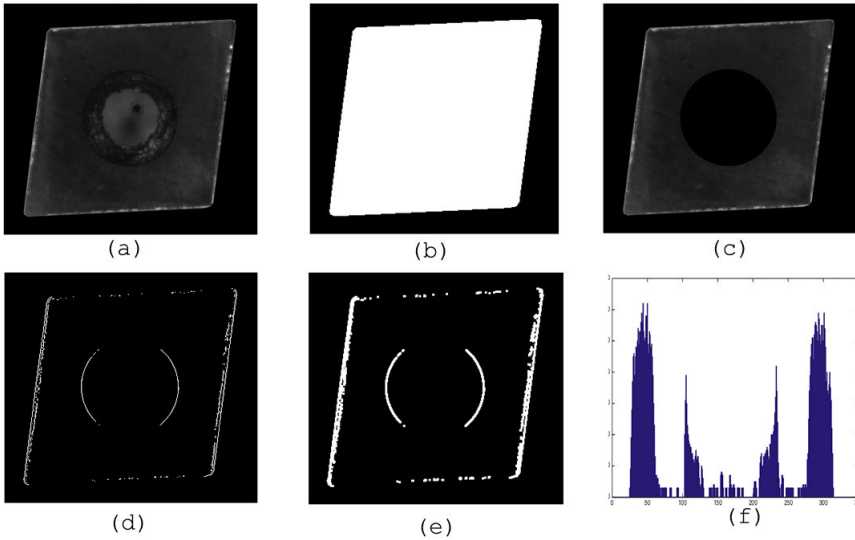


Figure 3.6: (a) Masked image of an insert. (b) Binary image after applying a threshold of 0.01. (c) Image with the center circle removed. (d) Image after sobel filter convolution. (e) Image after applying morphological operations (dilatation, opening). (f) Vertical projection of (e) to identify the edges.

First of all, the central portion of the insert was removed (Figure 3.6 (c)), masking out a circular region, so we first needed to locate its center and obtain the length of its radius. To determine the circle center, the image was binarized in order to remove the background, using a threshold of 0.01 - obtained empirically -, so the pixels of the insert area were set to 1 and the background to 0 (Figure 3.6 (b)).

The centroid of this area is considered as the center of the central circle (x_c, y_c) , where x_c and y_c are the mean value of all the x and y coordinate pixels with value 1, respectively. The radius of the circle R_c is considered to be $R_c = D/4.92$, where D is the length of the major diagonal of the insert and the ratio value, 4.92, was obtained taking into account previous knowledge of the size of the central circle respect the insert, which is the same in all the samples.

Thereafter, four regions with a width of one sixth of the insert and the same height, which contains the corresponding four cutting edges of the insert were extracted. We only explain how the west edge is extracted; the other three edges -north, east and south- were extracted using the same procedure rotating the original image 90° , 180° and 270° , respectively. Using the gray-scale image resulting from the previous step, a vertical Sobel filter with a kernel of 3×3 was applied to detect the inserts contours (Figure 3.6 (d)). These contours were then dilated and opened using a structuring element with square shape of size 3 (Figure 3.6 (e)), and vertically projected onto the horizontal axis (Figure 3.6 (f)). The first non-zero element in this projection indicates the x coordinate where the cutting edge starts. Then, the image of the region was cropped from the starting x position with a width of 100 pixels and the same height as the original image. A parametric margin of 25 pixels was added to each side in order to increase the crop area, because experimental tests showed that some inserts lost some edge pixels due to too tight crops.

Since inserts are rhomboid-shaped and their cutting edges are not aligned with the vertical and horizontal axis, the edges images were rotated to make the cutting edges to be aligned with the horizontal axis. To carry out this operation, a horizontal Sobel filter (3×3 kernel) followed by a morphological dilation was applied with the same structuring element used in the previous step. The resulting image was filtered to remove smaller objects leaving a binary image containing only the cutting edge. Finally, the orientation of the major axis of an ellipse that contains the edge was estimated, so the edge was rotated to compensate this orientation and leave it in horizontal position. Some examples of the cropping of cutting edges can be seen in Figure 3.7. Once we have obtained the regions with the edge information, the

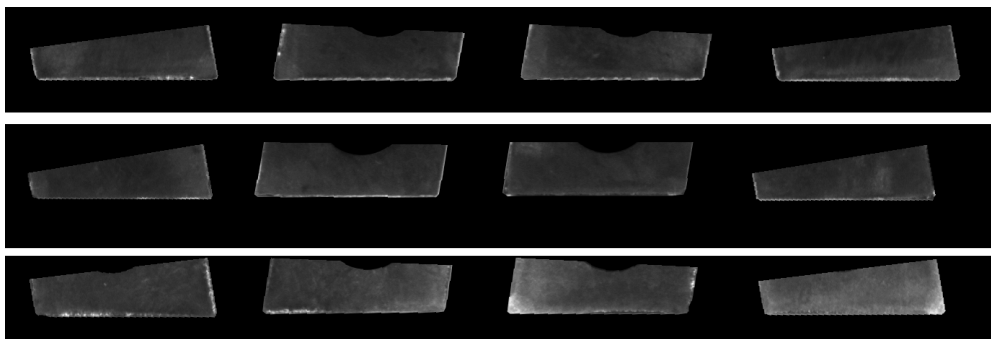


Figure 3.7: Rotated crops example.

next step consists of extracting the wear within each image. In order to calibrate the classification model, it is more important a correct wear shape segmentation than an

automatic one due to the need of having accurate shape information, so a manual ground truth of the wear region was segmented. At the end of this process, we got a dataset composed by 212 binary images. An example of the edges dataset is shown in Figure 3.8.

Figure 3.8: This is an example of the obtained dataset for the images shown in Figure 3.3. Each of the previous insert gray-scale images with masked background follow a preprocessing step that yields four new images, one for each cutting edge in horizontal position.

3.3. Insert regions datasets

The region dataset was created following the same process described in the previous section. Cropped images are usually formed by two different types of wear regions: complete and incomplete regions. The complete regions correspond with the horizontal cutting edge in the crop, and the incomplete regions represent the edges that converge in the complete one. To determine the wear of the complete edge by itself and also how it is influenced by the incomplete edges, we divided the whole Insert dataset into two: (a) Insert-C subset with the complete edges that correspond with the horizontal wear in the cropped images, and (b) Insert-I subset with the incomplete ones which orientations are vertical in the images. An example is shown in Figure 3.9.

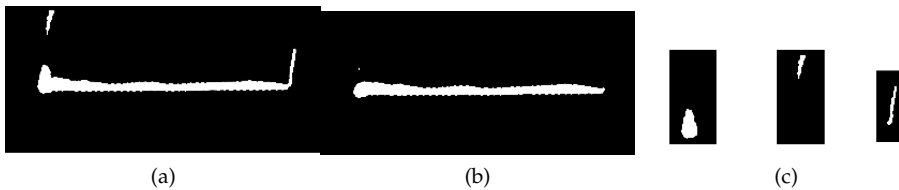


Figure 3.9: (a) The image containing both types of edges: The complete edge placed in horizontal position and the imcompletes edges in vertical position. (b) The complete edge and (c) Incomplete edges from vertical edges of (a).

An expert classified visually all the wear regions included in the dataset to generate the ground truth, carrying out two types of ratings depending on how many wear classes were taken into account: two (i.e., Low wear (L) with 260 images

and High wear (H) with 313 images) or three (i.e., Low wear (L) with 126 images, Medium wear (M) with 260 images and High wear (H) with 187 images). The expert carried out the labelling process by means of a visual assessment, relying on his previous knowledge and experience. This task is not easy as it depends on many parameters like the size of the wear area, its location or how deep it is.

In Tables 3.2 and 3.3 we can see the mean area (in mm^2) of the wear classification made by the expert and their standard deviation. The measures were carried out for the three datasets and distinguish between two and three wear classes. As we can see, the mean area is higher for the inserts categorized by the expert as high wear. This information resulted very interesting leading us to perform a classification of the wear based on these parameters.

Table 3.2: Mean area of the wear for the L-H labelling and their standard deviation (mm^2)

	Insert-C	Insert-I
L	27.25 ± 9.66	4.88 ± 2.06
H	48.55 ± 20.29	12.05 ± 6.43

Table 3.3: Mean area of the wear for the L-M-H labelling and their standard deviation (mm^2)

	Insert-C	Insert-I
L	21.64 ± 7.53	3.55 ± 1.47
M	34.72 ± 9.44	7.43 ± 2.24
H	51.76 ± 21.63	14.91 ± 7.17

3.4. Insert High resolution dataset for texture description

There was no publicly available datasets of milling processes with the enough quality and number of images to be used for texture description methods. For that reason, we created a new dataset composed of 254 images of an edge profile cutting head. The head tool has a cylindrical shape and it contains 30 inserts arranged in 6 groups of 5 inserts diagonally located along the axial direction of the tool perimeter. We have used a monochrome Genie M1280 1/3" camera with an active resolution of 2592×1944 which mounts an AZURE-2514MM lens with a focal length of 25mm. This camera is managed by using a Raspberry Pi which can be easily integrated in every manufacturing system due to its small size. In

order to achieve better illumination conditions, we have used three bars of LED (BDBL-R(IR)82/16H) which provide higher contrast in the edges of the inserts. It is important to notice that in that point, this machine tool does not use oils, lubricants or other kind of substances that can lead to low quality images. An example of the images obtained automatically by the capturing system can be shown in Figure 3.10



Figure 3.10: Example of the captured images.

After the image acquisition step, an automatic process to extract the cutting edge was carried out following the next steps: First of all, the screws located in the middle of the inserts were found using the Circular Hough Transform (CHT) to detect circles of radius between 40 and 80 pixels. This size is fixed due to the previous knowledge of the screw size for images of 2592×1944 . The next step consists of applying a Canny filter to detect the edges of the inserts. After that, we detect the vertical lines thanks to the standard Hough transform (SHT). Finally, we segmented the image taking into account the first vertical line located in the left side of the image, which corresponds with the cutting edge. The final result can be shown in the Figure 3.11. At the end, we have obtained 577 images of the cutting edge (301 intact edges and 276 damaged ones).

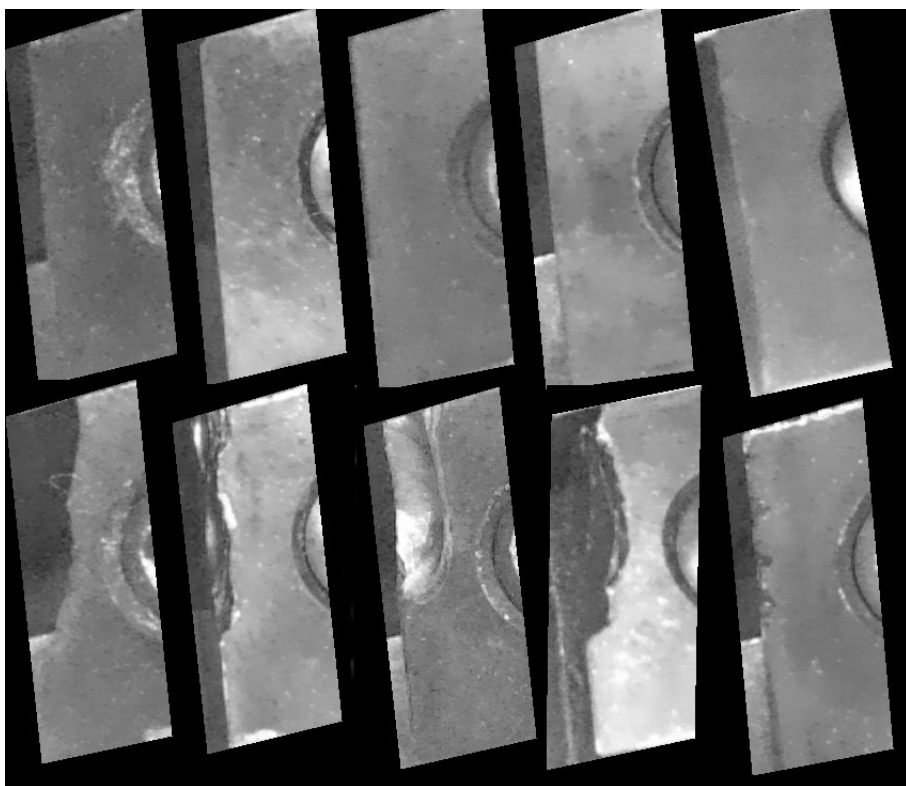


Figure 3.11: Cutting edge images. In the first row, intact edges are shown. In the second one, edges with different kind of wear are displayed.

Chapter 4

Contour based descriptors

In the last decades, the number of digital images has grown exponentially. For this reason the demand of techniques for indexing, storing and retrieving these digital images has become more and more necessary.

A lot of methods describing shapes have been developed over the years as we can see in the state of the art presented in Chapter 2. Images are usually described by features based on contour, color, shape or texture. Some of these features are also an essential factor for our particular problem of tool wear monitoring.

Two of the key factors in an image retrieval system are its accuracy and its computational efficiency. Image retrieval using shape information has received some attention by the image processing and pattern recognition community, focusing on improving any, or even both, of these factors.

In this chapter, two new methods based on contour information are proposed. Both of them are carried out in order to improve the accuracy and the computational cost of CPDH method which was developed by Shu et al. in 2011 (Shu and Wu, 2011b). In the next section we are going to introduce the original CPDH method and present our two new proposals.

4.1. Methodology

4.1.1. CPDH

The original method, Contour Points Distribution Histogram (CPDH), was proposed by Shu et al. (Shu and Wu, 2011b). It is based in analyzing the contour of an image and describe it taking into account the number of boundary points which belongs to a determinated region as well as its radius and angle. The first step is to obtain the points representing the contour of the image under polar coordinates. The contour points are used to describe the image and they can be obtained by different methods. In this case, they used the standard Canny operator to detect the object boundary (see (Belongie et al., 2002)). Once the points are extracted, the centroid is set as the origin. After that, they construct the minimum circumscribed circumference and the region defined by that circumference is divided into several bins using some concentric circumferences and equal interval angles. The final

step is to construct the CPDH descriptor taking into account the number of points belonging to each division, and the angle and radius of each division.

The main problem of CPDH, was that the authors classified using EMD distance, a very slow and time consuming method to make the descriptor invariant to rotation. Our descriptors can be classified using k-Nearest Neighbor (kNN) with classical metrics such as Euclidean, Intersect, Hamming, Cityblock,... distances which are less time consuming than EMD.

The full process to create our proposals based on CPDH is explained in the following sections.

4.1.2. CPDH36R

For creating CPDH36R, we also used the method proposed by Belongie et al. (Belongie et al., 2002), obtaining a contour like the one shown in Figure 4.1.

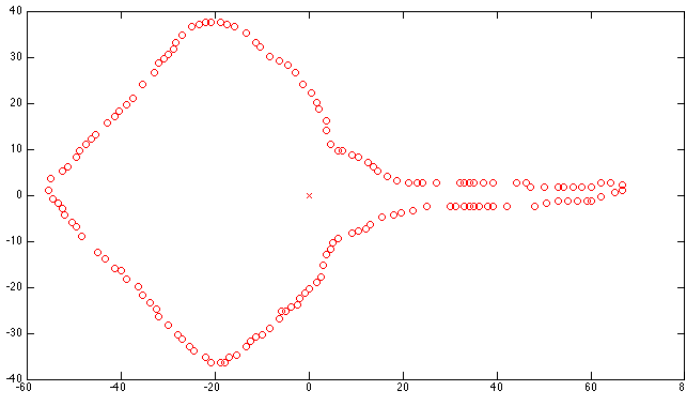


Figure 4.1: Contour points extraction using Belongie method.

This contour can be represented as a points collection as

$$P = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\} \quad (4.1)$$

$$/(x_i, y_i) \in \mathbb{R}^2$$

where m denotes the number of points on the contour. After that, the centroid of the previous points is obtained using the expression 4.2 in order to obtain the minimum shape circumscribed circumference.

$$(x_c, y_c) = \frac{\sum_{i=1}^m (x_i, y_i)}{m} \quad (4.2)$$

The centroid is set as the origin and the points collection, P , is translated into polar coordinates using the equation 4.3.

$$P = \{(\rho_1, \alpha_1), (\rho_2, \alpha_2), \dots, (\rho_m, \alpha_m)\} \in \mathbb{R}^2 \quad (4.3)$$

where

$$\rho_i = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2} \quad (4.4)$$

and

$$\alpha_i = \begin{cases} \arctan\left(\frac{y_i - y_c}{x_i - x_c}\right) & \text{if } x > 0, y \geq 0 \\ \arctan\left(\frac{y_i - y_c}{x_i - x_c}\right) + 2\pi & \text{if } x > 0, y < 0 \\ \arctan\left(\frac{y_i - y_c}{x_i - x_c}\right) + \pi & \text{if } x < 0 \\ \arctan\left(\frac{\pi}{2}\right) & \text{if } x = 0, y > 0 \\ \arctan\left(\frac{3\pi}{2}\right) & \text{if } x = 0, y < 0 \end{cases} \quad (4.5)$$

Therefore, the circumscribed circumference has as centre the centroid $C(x_c, y_c)$ and radius $\rho_{max} = \max(\rho_i), i = 1, 2, \dots, m$.

Then, the circle defined by the circumscribed circumference is divided into several bins partitioning its area into $u \times v$ bins being u the parts in which ρ_{max} is divided and v the sectors in which the circle is divided. The result of this process is shown in Figure 4.2.

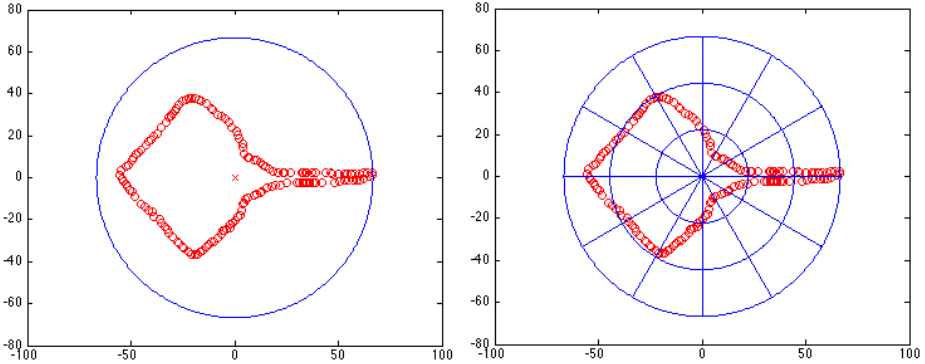


Figure 4.2: Circumference circumscribed to the image contour (left) and circle partition in 36 bins ($v=12, u=3$) (right).

After these previous steps, we can build the descriptor by counting the number of points belonging to each bin. The original CPDH is composed by a triplet for each bin $H_i = (\rho_i, \theta_i, n_i)$ where ρ_i denotes the radius of the concentric circumferences, θ_i denotes the angle space and n_i denotes the number of points located in the bin r_i .

The authors of CPDH used the EMD (Earth mover Distance) classifier in order to make their descriptor invariant to rotation. The main disadvantage of EMD is its

slowness, thus one of our motivations in developing Contour Points Distribution Histogram Radius using 36 features (CPDH36R) was the necessity of making a new descriptor containing just the most relevant features, and for this reason obtaining even better hit rates than with the base-line method but being much more efficient in terms of computational cost.

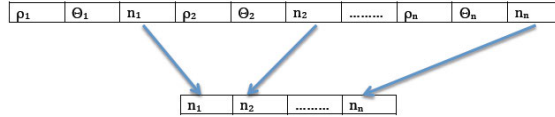


Figure 4.3: First step to construct CPDH36R: 36SEQ extraction.

The proposed CPDH36R descriptor uses the first and third element of each of the previous triplets, that are the radius and number of points in the bin respectively. Radius are normalized following equation 4.6

$$r_p = \frac{\rho_{1..u}}{\rho_{max}} \quad (4.6)$$

Where ρ_{max} is the biggest radius and $p \in [1..36]$. CPDH36R is created multiplying the number of points of each bin (Figure 4.3) by its corresponding normalized radius (first element of the triplet), that contains just the most relevant features for classification given that the second element is constant to all the images because the number of circumference divisions is always the same as can be seen in Figure 4.4. As it will be shown in section 4.2.2, this proposal combined with kNN classifier outperforms the original CPDH descriptor with EMD in terms of accuracy and computational cost.

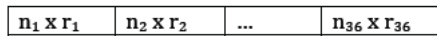


Figure 4.4: CPDH36R constructed by multiplying the number of points of each bin by its corresponding normalized radius

4.1.3. RCPDH

In this case, we also begin using the method proposed by Belongie et al. (Belongie et al., 2002) and follow the same steps explained in the previous section.

Basically, Rotational Invariant Contour Points Distribution Histogram (RCPDH) consists of making the CPDH descriptor intrinsically invariant and to avoid the use of EMD distance because of its time cost.

Once we have obtained the descriptor CPDH original as is described in section 4.1.1 using equations 4.2 and 4.3, the third element of each triplet (number of points in the bin) is extracted to create the 36SEQ as we explained in the previous section (See Figure 4.3).

Then, we search for the minimum module sequence in the 36SEQ descriptor shifting the vector along all the 36 positions following the expression 4.7. See Figure 4.5.

$$RCPDH = \text{cat}(CPDH(\rho_i, \theta_i, n_i : \rho_{u \times v}, \theta_{u \times v}, n_{u \times v}), CPDH(\rho_1, \theta_1, n_1 : \rho_{i-1}, \theta_{i-1}, n_{i-1})) \tag{4.7}$$

where i satisfies equation 4.8 and $\text{cat}(a, b)$ is the concatenate operator between a and b .

$$\text{arg}_i \min \{ \text{cat}(CPDH36(i : u \times v), CPDH36(1 : i - 1)) \forall i \in (1..u \times v) \} \tag{4.8}$$

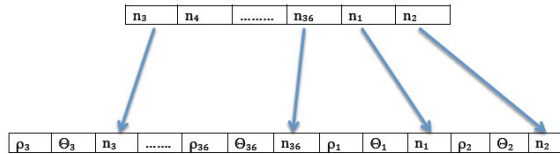


Figure 4.5: RCPDH descriptor.

Employing this new descriptor, it is not necessary the use of a classifier to resolve the rotation invariance problem, so we can utilize several classifiers solving the problem of the cost of time derived from EMD method. As shown in section 4.2.2, this proposal combined with kNN as well as CPDH36R also outperforms the original CPDH descriptor with EMD.

4.2. Shape recognition

4.2.1. Experimental setup

Three different kNN classifiers were used for evaluating shape recognition. Original kNN, mean distance kNN in which the average distance between the query

and the samples is calculated and the class is assigned taking into account this distance, and finally weighted kNN in which we take into account the neighbours depending on its assigned weight.

Mean Distance kNN

The main idea of using mean distance kNN is to assign the query sample to the class whose mean distance to the query sample is smaller, instead of assigning it to the most represented class. In Figure 4.6, we can see that although the most represented class is the X class, four samples against three, the query sample belongs to '-' class because the mean distance is smaller.

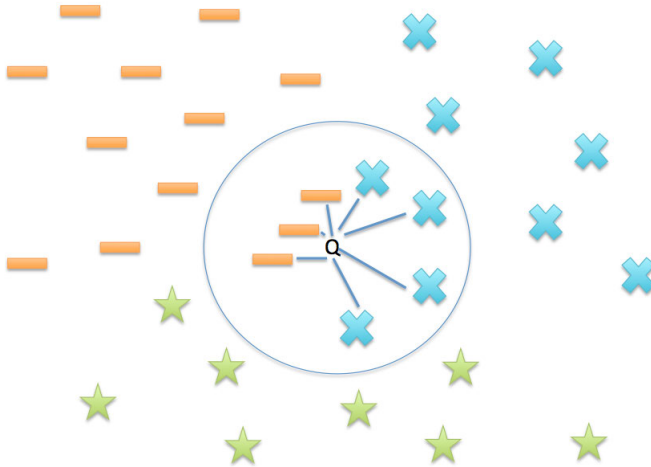


Figure 4.6: kNN mean distance example taking $k=7$.

Weighted kNN

In weighted kNN, a weight is assigned to each sample. This weight w_i is the difference between the farthest sample from the query and the proper distance from the query sample. See equation 4.9

$$w_i = |d_{max} - d_i| \quad (4.9)$$

Each class is weighted adding the weight of each sample of the class. Finally, the query sample belongs to the heavier class.

4.2.2. Results

Our proposals were classified using kNN, kNN Mean and Weighed kNN with intersect and cityblock distances and they were tested on four datasets: kimia25, kimia99, MPEG7 and MPEG2. Classification results for each dataset using the original kNN with k=1 are shown in tables 4.1 and 4.2. The case of Kimia25 is the only one in which the best result was achieved using EMD because of the small number of samples, (see table 4.1) obtaining a 90% of succes rate against the 76% for RCPDH and 84% for CPDH36R method. In table 4.1, we can see that results are better testing the descriptors against kimia99 dataset than against kimia25 because of the small number of samples and in most cases they are slightly better with our proposals CPDH36R and RCPDH than with the original CPDH. Seeing results in table 4.1 we can conclude that our proposal CPDH36R is the best one, so we tested it using the more challenging datasets:

In table 4.2 we can also see the efficiency of our CPDH36R descriptor. Using MPEG7 dataset, once again, our proposal CPDH36R outperforms the state of the art method (CPDH) with all the distances. Using MPEG2 dataset our method also outperforms the original one classifying with all the distances achieving a 91.25% of success rate using cityblock distance. Although on simpler dataset as Kimia25, the original CPDH with EMD offered better results, in the more complex dataset, CPDH36R has obtained a significant improvement of the hit rate against the original method with either EMD or kNN.

It is important to notice than for MPEG7 experiment, the combination of CPDH and EMD could not be tested due to the incapability of evaluating it with our computers due to it high computational cost and the size of the dataset (1400 images).

Table 4.1: Kimia25 and Kimia99 classification using CPDH original RCPDH and CPDH36R methods with k=1.

Classifier	Kimia25			Kimia99		
	CPDH	RCPDH	CPDH36R	CPDH	RCPDH	CPDH36R
kNN Intersect	60%	64%	80%	83.83%	89.89%	93.93%
kNN Cityblock	68%	76%	84%	91.91%	91.91%	94.94%
EMD	90%	-	-	86.41%	-	-

In Figures 4.7 and 4.8 we can see the results with our best proposal and the original CPDH classifier, kNN mean distance and weighted kNN for different values of k and Cityblock distance. As we can see in all the figures, using original kNN, the recognition hit rate decreases a lot for higher values of k. This problem is solved using kNN mean distance or weighted kNN in which the hit rate variations

Table 4.2: MPEG7 and MPEG2 classification using CPDH original and CPDH36R methods with $k=1$.

Classifier	MPEG7		MPEG2	
	CPDH	CPDH36R	CPDH	CPDH36R
kNN Intersect	50.36%	82.07%	67.50%	91.25%
kNN Cityblock	59.14%	82.14%	74.75%	91.25%
EMD	-	-	44.25%	-

is less than using the original kNN. In Figure 4.7, using Kimia25 dataset, we can see that the original kNN performance is really bad for higher values of k , obtaining more stable results using kNN mean distance or weighted kNN. In the rest of the datasets, Figure 4.7, Figure 4.8, it is shown that the most stable results are achieved using kNN mean distance.

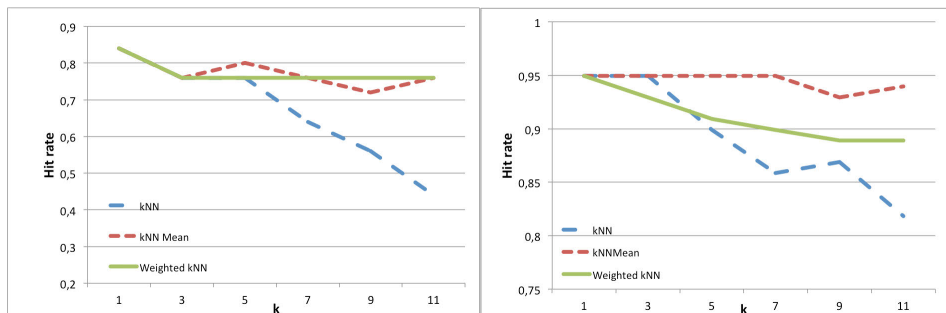


Figure 4.7: Hit rate of the different kNN methods applied on the CPDH36R descriptor using Kimia25 dataset (left) and Kimia99 dataset (right).

4.3. Conclusions

In this chapter, two new methods were proposed in order to retrieve objects based on their shapes: RCPDH and CPDH36R, both of them based on CPDH descriptor. The original method, CPDH, was classified using EMD distance, a very slow and time consuming method. Our descriptors can be classified using kNN with classical metrics such as Euclidean, Intersect, Hamming, Cityblock,... distances obtaining results in a few seconds while with EMD and these same datasets it takes days. The tests carried out using the original CPDH method and the classical metrics offered worse results than our proposals RCPDH and CPDH36R. As the CPDH36R

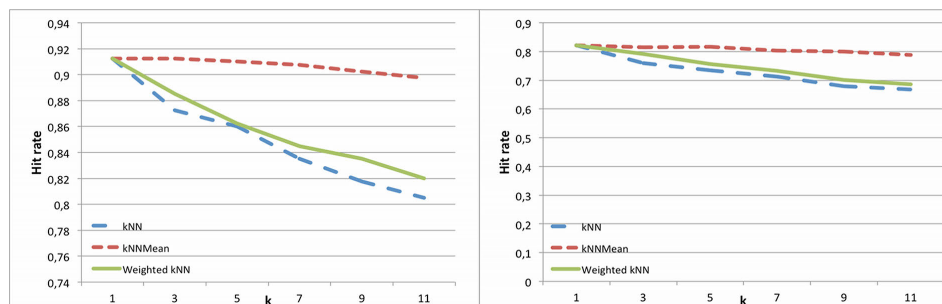


Figure 4.8: Hit rate of the different kNN methods applied on the CPDH36R descriptor using MPEG2 (left) and MPEG7 (right) datasets.

descriptor achieved better performance than RCPDH in simple datasets, it was tested using more challenging shape datasets. For this reason, CPDH36R was tested using four datasets: kimia25, kimia99, MPEG7 and MPEG2 achieving not only the highest differences in the more challenging one, the MPEG7, but also obtaining great performance in the MPEG2 (91.25%), and Kimia99 (94.94%) datasets.

Chapter 5

Moments and orientations based descriptors

Over the years, the computer vision research community has successfully developed new methods which use shape information in order to describe the images (Mukanova et al., 2014; Laiche et al., 2014; Shu and Wu, 2011a). Shape description methods have already been applied in the tool wear monitoring field to ensure the optimal replacement of the tool in lathe processes. In (Barreiro et al., 2008) the use of Hu and Legendre moments was proposed to deal with this problem. Our problem is quite different from that of (Barreiro et al., 2008) in the sense that we determine the tool wear in edge profile milling processes instead of lathe processes. In milling, the wear shape is distributed uniformly along the insert while in lathe processes the wear occurs on the same area concavely. Also, most works that deal with milling focus on end milling or face milling. Thus, monitoring tool wear in edge profile milling processes states a more challenging task since the insert distribution is not typical and the head geometry is more complex. Other description method based on moments is described in (Antić et al., 2018) where the authors consider the module of the Short Term Discrete Fourier Transform (STDFT) spectra obtained from the particular vibration sensors signal utterance as the 2D textured image. This is done by identifying the time scale of STDFT as the first dimension, and the frequency scale as the second dimension of the particular textured image.

In order to characterize the tool wear, in Chapter 4, we proposed contour descriptors. This kind of description has limitations because the contour does not show enough information about the level of wear. Therefore, we moved to moments and orientations shape descriptions with the aim of outperforming our previous results. In this chapter we propose to use shape descriptors computed from the tool image. In particular, the combination of global and local shape descriptors is considered to be quite reliable and has been extensively applied lately (Singh and Pooja, 2012; Anuar et al., 2013). On the one hand, local descriptors are used to characterize small particular regions of the image object. On the other hand, global descriptors provide comprehensive information of the whole image. However, they are not as robust as local ones because they are not sensitive to changes in small parts of the image. One example is the work carried out by Singh et al. (Singh and Pooja, 2012), that combines the Angular Radial Transform as global descriptor and

Polar Hough Transform as local one.

In this chapter, two novel shape descriptors based on moments and orientations of the contour, which also combine global and local descriptors are proposed. Both of them are based on the method developed by Anuar et al called ZMEG (Anuar et al., 2013). In the next sections we are going to introduce the original ZMEG method and present the two new proposals: aZIBO and B-ORCHIZ. Exhaustive experiments have been carried out that demonstrated the high performance of both methods not only for tool wear monitoring but also in more general problems.

5.1. Methodology

5.1.1. ZMEG

ZMEG (Anuar et al., 2013) is a shape descriptor which combines global and local descriptors on binary images. It makes use of the Zernike moments as the global shape descriptor (specifically the module of the first 36 coefficients, up to the tenth order). It is well known that the module of the Zernike moments of an image f has the property of being invariant to rotation.

The local shape descriptor is called EGCM (edge gradient co-occurrence matrix) and it is obtained as follows: First of all, the boundary points of the image are obtained. In this case, the edge detector proposed by Belongie et al. (Belongie et al., 2002) which yields a uniformly spaced sample of the points was used. The result of this process is shown in Figure 5.1.

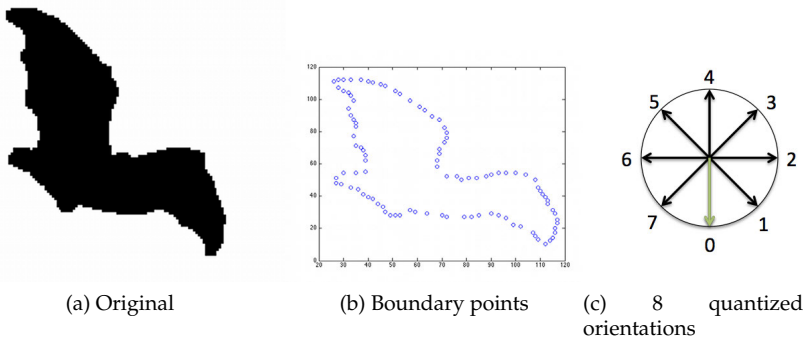


Figure 5.1: Steps followed to obtaining the eight quantized orientations.

The set of boundary points is represented as:

$$P = \{(x_i, y_i) \in \mathbb{R}^2 : 1 \leq i \leq m\},$$

where m is the number of boundary points.

In the second step, the gradient of each boundary point in P is calculated by means of Eq.(5.1).

$$O(x, y) = \arctan \left(\frac{I(x + 1, y) - I(x - 1, y)}{I(x, y + 1) - I(x, y - 1)} \right), \quad (5.1)$$

where x and y are the image coordinates.

These gradient orientations are discretized into the 8 values shown in Figure 5.1(c).

In the third step, the EGCM matrix is built taking into account only the points among its eight neighbors which are part of the boundary. Figure 5.2 depicts the EGCM construction for a boundary point.

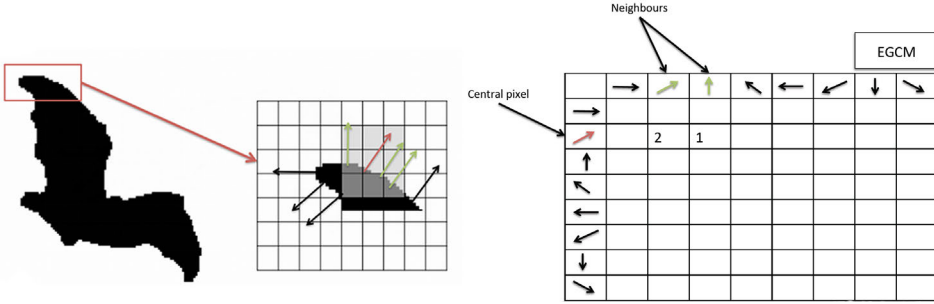


Figure 5.2: Example of the orientations taken into account for the boundary point shown in red for the composition of the EGCM (left). Example of the composition of the edge gradient co-occurrence matrix taking into account the boundary point shown in this figure (right).

Finally, the eight rows of the EGCM matrix are concatenated yielding a 64-element vector. In the work developed by Anuar et al. (Anuar et al., 2013), these descriptor values were normalized in the range of 0 to 1.

5.1.2. aZIBO

Our first novel proposal was aZIBO descriptor. The global shape descriptor of aZIBO follows the baseline method in ZMEG, but using the images with intensity values 0 and 255 (not binary ones) and resizes them to 128×128 to give more information to the Zernike moments coefficients thanks to the interpolation process which generates a grey-scale image, not just a binary one.

The local descriptor EGCM is made rotationally invariant. To do this, the pair of boundary points b_1 and b_2 whose Euclidean distance is the longest is found, and then the angle of vector $\vec{b} = b_1 \vec{b}_2$ is calculated by means of Eq.(5.2).

$$\alpha_{\bar{b}} = \arctan \left(\frac{b_2(2) - b_1(2)}{b_2(1) - b_1(1)} \right), \quad (5.2)$$

In case there are more than one pair of points with such distance, it is taken the pair (b_1, b_2) for which the segment joining them, $\overline{b_1 b_2}$, is the nearest to the centroid of the shape. The point b_1 or b_2 with more boundary points in its 21×21 neighborhood is selected as origin. The size of the neighborhood has been determined experimentally. The gradient orientation of the origin, ϕ_d , is placed at the first row and column in the edge gradient co-occurrence matrix and the rest of orientations are shifted following their original position. Therefore, if EGCM is $[\phi_1, \phi_2, \dots, \phi_d, \dots, \phi_8]$, then IEGCM (the rotationally invariant version of EGCM) is $[\phi_d, \dots, \phi_8, \phi_1, \dots, \phi_{d-1}]$.

An example of this process is shown in Figure 5.3, where the same matrix IEGCM is obtained for the same image in two different orientations.

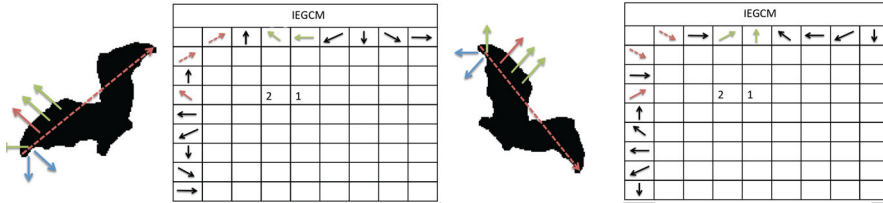


Figure 5.3: Example of the invariance to rotation of the IEGCM. Red dotted line represents the orientation between the furthest points in the boundary. The red arrow corresponds to the gradient orientation of the point whose IEGCM is calculated. The green arrows are taken into account (because they are located inside the neighborhood) whereas the blue arrows represent points that are not in the influence area of the selected one.

Finally, the eight rows of the IEGCM are concatenated yielding a 64-element vector. Unlike the work in (Anuar et al., 2013) where these descriptor values were normalized in the range of 0 to 1, aZIBO preserves the original values in order to obtain as much information as possible.

5.1.3. B-ORCHIZ

In this section, we present a novel shape descriptor as an extension of ZMEG and aZIBO attempting to improve its reliability and robustness. BORCHIZ improves the previous method (aZIBO) by adding more information about the orientations of the contour points of the wear region. This allows us, on the one hand, to have a global description based on moments, and on the other hand, a complete local description based on orientations, being both descriptions invariant to rotation.

B-ORCHIZ uses the Invariant Boundary Descriptor (IBD) - a fusion of the IEGCM and the Boundary Orientations Chain (BOC) - as local shape descriptor. In addition, it achieves rotational invariance in the global descriptor using the module of Zernike moments up to the tenth order in the same way as aZIBO does.

The first component of the local descriptor, IEGCM, was described in the previous section. Next, we present the second local descriptor, the Boundary Orientations Chain (BOC). Boundary Orientations Chain is a vector whose elements are the orientations of the m boundary points. As the other descriptors composing B-ORCHIZ, BOC must be also invariant. In order to achieve invariance to rotation, the first element must be determined. The BOC extraction process is the following:

1. The first step is to find the two furthest contour points b_1 and b_2 , which is done following the same process as it was done to find the first orientation in the IEGCM.
2. The selected point b_1 or b_2 will be the one with more boundary points in its 21×21 neighborhood.
3. Thereafter, the orientation values are normalized to achieve invariance to rotation so the first orientation, whatever it is, is considered as the value 0 in the chain and the rest are shifted according to it, as is shown in Figure 5.4.

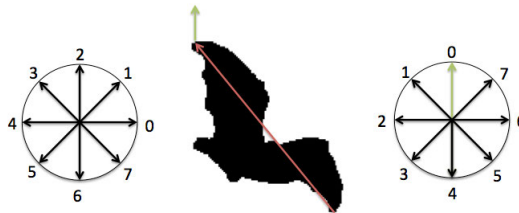


Figure 5.4: Example of the Boundary Orientations Chain (BOC) rotation invariance. The red arrow represents the maximal distance between points in the contour of the image. The green arrow is the orientation of the first point of the chain. The values assigned to each orientation are shifted locating the green arrow in the top of the diagram (value 0).

After that, the rotationally invariant descriptor is the vector of m elements representing the orientation of each boundary point. In this work, the local descriptor IBD is the concatenation of the previous IEGCM and the BOC. After all this process, the global and local descriptors are concatenated.

5.2. Shape recognition application

5.2.1. Experimental setup

In this section, we evaluated the performance of B-ORCHIZ, aZIBO and ZMEG on three of the most commonly used datasets in shape recognition: Kimia-99, MPEG-7 and MPEG-2 datasets. These images were classified using the k-Nearest Neighbor classifier (k-NN) with Cosine, Euclidean, Intersect, Chi-square and Cityblock distances. We have selected this classifier because it is the most used in all the experiments that try to evaluate descriptors using these datasets (Kimia and MPEG7).

As it was pointed out, B-ORCHIZ is a combination of global (the 36 modules of Zernike moments) and local descriptors (the invariant boundary descriptor, or IBD, which is a fusion of IEGCM and BOC). The high difference of magnitude of both local and global parts is very remarkable. Therefore, it is necessary to carry out a normalization or a weighting to ensure that the classification step is not biased by such differences.

In our case, we weight both parts using (5.3).

$$B_{ORCHIZ} = Z \cdot w_g + IBD \cdot w_l, \quad (5.3)$$

where Z stands for the vector of the 36 modules of Zernike moments. The values of the weights have been chosen experimentally, fixing $w_g = 1$ and $w_l = 5$.

5.2.2. Results

Table 5.1 reports the classification accuracy achieved with the k-NN classifier using the shape descriptors B-ORCHIZ, aZIBO and ZMEG. We can see that B-ORCHIZ outperforms the other methods on the three datasets regardless the distance metric used.

As it is shown in the first part of Table 5.1, the best accuracy for the Kimia dataset, 91.92%, is obtained by B-ORCHIZ using the cosine distance, while ZMEG and aZIBO gets a hit rate of 87.88%.

Regarding the results on MPEG-2 dataset (see table 5.1, second part), B-ORCHIZ achieves 93.75% of accuracy using cosine distance against 79.50% and 79.00% achieved by aZIBO (euclidean) and ZMEG (city block), respectively.

It is remarkable that in all the experiments the best results were obtained taking into account just the nearest neighbor (i.e., $k = 1$), whereas the classification performance decreases as the size of the neighborhood increases in almost all tests, possibly due to the great similarity between the forms of the different classes.

With respect to the results on the most challenging of the assessed datasets, MPEG-7, shown in the third part of Table 5.1, once again B-ORCHIZ outperforms

the other descriptors for whatever distance, what points out its improvements with respect to the previous proposals. In this case, B-ORCHIZ obtained an accuracy of 89.50%, while aZIBO and ZMEG achieve 77.07% and 76.57%, respectively.

The full rotational invariance of B-ORCHIZ can perfectly explain the highest rates in all the experiments. Furthermore, thanks to not only Zernike moments and IEGCM but also to the Boundary Orientation Chain (BOC), the description of complex shapes has been reinforced. With this new features, B-ORCHIZ becomes in the descriptor with a higher discriminant power of all the three.

Table 5.1: Classification accuracy in % of our proposal B-ORCHIZ against aZIBO and ZMEG using kNN (with different distances) in the Kimia-99 (first), MPEG-2 (second) and MPEG-7 (third) database.

	kNN Cosine	kNN Euclidean	kNN Intersect	kNN Chi-square	kNN Cityblock
B-ORCHIZ	91.92	90.91	90.91	89.90	90.91
aZIBO	85.86	86.87	86.87	86.87	87.88
ZMEG	87.88	85.86	83.84	72.73	83.84
B-ORCHIZ	93.75	93.00	92.25	88.50	92.75
aZIBO	72.25	79.50	75.75	79.25	79.00
ZMEG	75.25	79.00	52.75	61.25	79.00
B-ORCHIZ	89.50	88.29	87.29	84.36	88.21
aZIBO	75.00	75.79	73.29	76.50	77.07
ZMEG	75.93	76.29	38.36	56.79	76.57

The precision-recall curves of the classifications using B-ORCHIZ and the other assessed descriptors with different values of k are shown in Figure 5.5. As we can see clearly, B-ORCHIZ descriptor outperforms the other methods in the three classifications scenarios.

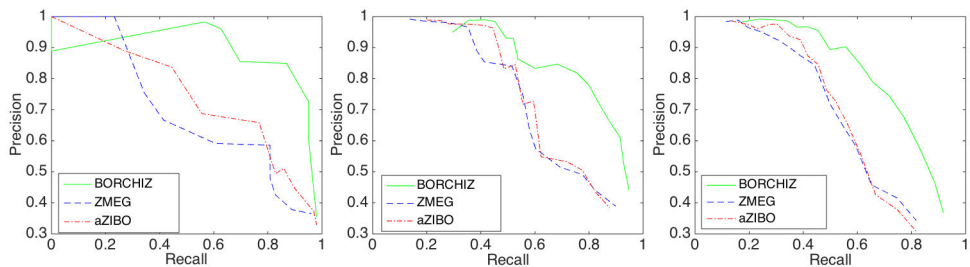


Figure 5.5: Precision-Recall curves of the classification using Kimia-99, MPEG-2 and MPEG-7 dataset with B-ORCHIZ, aZIBO and ZMEG.

These results obtained for the classical shape datasets, confirms the good performance of B-ORCHIZ and also aZIBO in comparison with the method proposed by (Anuar et al., 2013) for all kind of shape recognition processes.

5.3. Tool Wear monitoring

We evaluated our proposals in an image-based tool wear monitoring system for edge profile milling. Their efficiency in terms of the accuracy to determine the wear level was compared with ZMEG (Anuar et al., 2013), that also use local and global features. We used two datasets to test our methods. The insert edges dataset and the insert region dataset. This last one encompasses more information so additionally we carried out more tests such as a cluster analysis, a study of the representative image for each class and also a wear size evaluation.

5.3.1. Experimental setup

In the case of the Insert edge dataset, the k -Nearest Neighbor classifier was used with k equal to 1, 3, 5, 7, 9 and 11 and six distances: Cosine, Euclidean, Intersect, Chisquare, SqDist and Cityblock to classify the concatenation of the global and local shape descriptors explained in the methodology section. Two types of classification were applied: the first one considering three classes (low (L), medium (M) and high (H) wear) and the second one considering two classes (low (L) and high (H) wear).

In the Insert region dataset, we evaluated empirically the classification accuracy achieved with previously proposed shape descriptors and compared them with aZIBO and B-ORCHIZ.

In all cases, classification was performed using a Support Vector Machine (SVM) with linear, second order polynomial and intersection kernels. The Least Square algorithm was used as training method. The classification has been carried out using a stratified 10-fold cross validation. Therefore, the whole dataset (i.e. Insert-C, Insert-I or the complete dataset) is partitioned randomly into 10 equally sized subsets with the same distribution as the original dataset. Of the 10 subsets, 9 are used to train the model and the remaining one is used as the test set. The cross validation is repeated 10 times using a different subset each time as test set. The 10 results from the 10 folds are averaged to provide the final result in terms of accuracy, which is calculated as the proportion of true results (both true positives and true negatives) among the total number of cases examined.

5.3.2. Results on the insert edge dataset

As aforementioned, the classification was carried out using six different distance metrics. In Tables 5.2 and 5.3 the best hit rates obtained in the classifications using three description methods (aZIBO, Hu and Flusser) are shown. Cosine, ChiSquare, SqDist and Cityblock represent always the best cases. As shown, aZIBO achieves better results than classical descriptors (Hu and Flusser) obtaining a hit rate up to

60.37% when using three wear classes and 81.13% when using two wear classes. In this experiment $k=9$ has been used.

Table 5.2: Inserts classification using aZIBO against classical Hu and Flusser descriptors with k equals to nine and using Cosine, ChiSquare, SqDist and Cityblock distances with 3 classes labeling.

	Cosine	ChiSquare	SqDist	Cityblock
aZIBO	58.96%	57.07%	60.37%	59.43%
Hu	55.66%	56.13%	59.91%	58.49%
Flusser	57.07%	58.96%	57.08%	57.08%

Table 5.3: Inserts classification using aZIBO against classical Hu and Flusser descriptors with k equals to nine and using Cosine, ChiSquare, SqDist and Cityblock distances with 2 classes labeling.

	Cosine	ChiSquare	SqDist	Cityblock
aZIBO	81.13%	80.66%	81.13%	81.13%
Hu	76.89%	79.72%	79.25%	80.19%
Flusser	79.25%	78.30%	75.47%	75.47%

Figure 5.6 shows with more detail all the results for aZIBO, Hu and Flusser, for all k values and the six distances.

As we can see, L-H classification always offers better results than L-M-H classification although it provides less information. Also, aZIBO outperforms classical descriptors in almost all cases.

5.3.3. Results on the insert region dataset

In this section we evaluated our proposals in an image-based tool wear monitoring system for edge profile milling. Also, a cluster analysis allows to study the shape descriptor capacity to cluster insert images in homogeneous groups. Furthermore, this cluster analysis provide prototype images for certain tool wear stages with the goal to visualize where and how the wear takes place and how the typical wear shape is. We also evaluated the classification accuracy obtained using the wear size (area in mm^2) but results demonstrated the fact that the area information is not enough to describe precisely the wear level. Techniques based on shape description appear to be a better option to assess the wear level of the inserts.

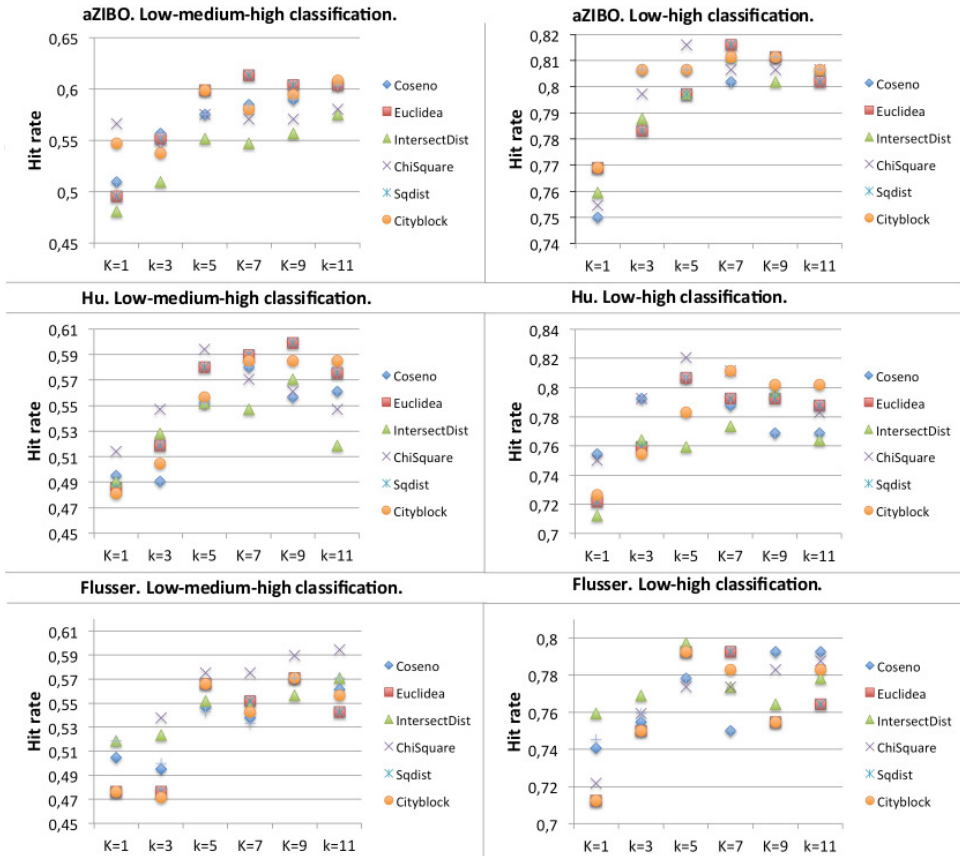


Figure 5.6: kNN classification for the six distances and different values of k using aZIBO, Hu and Flusser and two kinds of labeling datasets: L-M-H and L-H classes.

Descriptor evaluation for tool wear classification

In this section, we evaluated empirically the classification accuracy achieved with B-ORCHIZ, aZIBO and ZMEG in the insert regions dataset. Both subsets (Insert-C and Insert-I) were evaluated, first separately and later both together in order to study the possible influence of vertical edges into the main cutting edge.

Classification accuracies for each of the subsets are shown for the L-H and L-M-H labelling in Tables 5.4 and 5.5. Table 5.6 shows the accuracies using both subsets.

For the Insert-C subset (Table 5.4), B-ORCHIZ and aZIBO achieved an accuracy of 87.02% in the binary classification using the intersection kernel while ZMEG obtained just 85.58%. In the more challenging three-class classification, the performance of B-ORCHIZ was higher than the other two methods, 81.25%, in

contrast with the 76.92% and 76.44% of aZIBO and ZMEG, respectively. It is remarkable that in all the experiments, the intersection kernel has shown the best performance, improving the three-class classification in a 35.19% and a 60.95% with respect to the linear and the polynomial kernels, respectively.

Table 5.4: Classification Accuracy (in %) using SVM (linear, quadratic and intersection kernels) of the two and three-class scenarios for the Insert-C dataset.

Classifier	L-H			L-M-H		
	B-ORCHIZ	aZIBO	ZMEG	B-ORCHIZ	aZIBO	ZMEG
SVM-Linear	70.77	72.12	73.08	60.10	58.17	58.17
SVM-Quadratic	80.29	75.96	77.88	50.48	49.52	49.04
SVM-Intersection	87.02	87.02	85.58	81.25	76.92	76.44

In Figure 5.7 we can see some examples of misclassified images. The first column shows some images whose ground truth (ie., provided by the expert) is low wear and have been misclassified as having high wear, whereas in the second column the opposite case is depicted. As we can see, the three examples that actually have low wear are large, even wider than the third image with high wear. It is also noticeable that two of the images misclassified as low wear have a very small wear region, which can be easily confused with a low wear one. As the system works with binary images, it has no possibility to take into account the depth of the wear, which could be one of the reasons for the expert to label them in this way.

In the case of the Insert-I dataset (see Table 5.5), the behaviour of all the methods is very similar to the previous analysis. In the binary problem, B-ORCHIZ outperforms all the other methods achieving an accuracy of 88.46%: 4.21% and 5.91% higher than aZIBO and ZMEG respectively. In the L-M-H classification, ZMEG and aZIBO obtained hit rates of 79.42% and 82.14% respectively, while B-ORCHIZ achieved 82.69%. Once again, the best performance was achieved using the intersection kernel in all the tests.

Table 5.5: Classification Accuracy (in %) using SVM (linear, quadratic and intersection kernels) of the two and three-class scenarios for the Insert-I dataset.

Classifier	L-H			L-M-H		
	B-ORCHIZ	aZIBO	ZMEG	B-ORCHIZ	aZIBO	ZMEG
SVM-Linear	78.02	76.37	76.65	61.54	70.60	72.25
SVM-Quadratic	76.37	67.58	74.84	61.26	60.16	64.01
SVM-Intersection	88.46	84.89	83.52	82.69	82.14	79.40

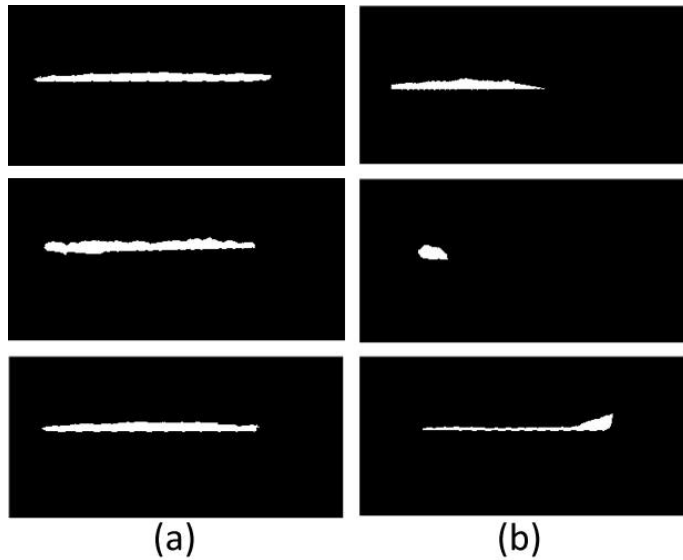


Figure 5.7: Examples of misclassified images: (a) low wear inserts misclassified as high wear and (b) high wear inserts misclassified as low wear.

The third experiment consisted in using the whole Insert dataset, which is more challenging due to the intra-class variation of the dataset. The aim was to assess the influence of the vertical edges on the final evaluation of the cutting edge. The experiments demonstrated (see Table 5.6) that in terms of classification accuracy the fusion of both subsets does not improve classification leading to similar results to the Insert-C dataset and slightly worse than the Insert-I dataset.

Table 5.6: Classification Accuracy (in %) using SVM (linear, quadratic and intersection kernels) of the two and three-class for the Insert dataset (full dataset).

Classifier	L-H			L-M-H		
	B-ORCHIZ	aZIBO	ZMEG	B-ORCHIZ	aZIBO	ZMEG
SVM-Linear	80.42	78.32	79.90	62.59	63.11	66.96
SVM-Quadratic	81.12	71.33	72.03	72.52	57.69	61.36
SVM-Intersection	87.06	84.44	83.74	80.24	78.85	75.87

The precision-recall curves of the classifications using B-ORCHIZ and the other assessed descriptors for Insert-C, Insert-I and the whole Insert dataset are shown in Figure 5.8. It can be seen that B-ORCHIZ leads to the best performance for the three scenarios considered.

Results show that combining the complete edge data with the incomplete edge data does not contribute to improve the classification accuracy. On the other hand, the incomplete edge data by itself allows to assess the wear level with a slightly higher accuracy compared to the one achieved with the complete edge data. These results can be explained because in the case of the incomplete edge, all the extracted information is close to the corners of the inserts, which usually are the places where the wear is shown sooner and stronger. For this reason, using just the information of the corners and avoiding the central part of the cutting edge, the precision of the system increases. In the Precision-Recall curves (Figure 5.8) we can also see how the precision decreases faster when both datasets are combined than when they are evaluated individually, remaining with accuracy equal to 1 for recall values higher than 0.7 in both cases.

As Fernandez-Robles et al. pointed out (Fernández-Robles et al., 2015a), the resting time of milling head tools lies between 5 and 30 minutes. Our experiments, ran on a computer with an i7 processor and 16GB RAM using MATLAB, showed that it takes about 60 seconds to describe the full insert dataset and less than 0.1 seconds to classify all the training data, so the implementation reaches real time performing on a production environment.

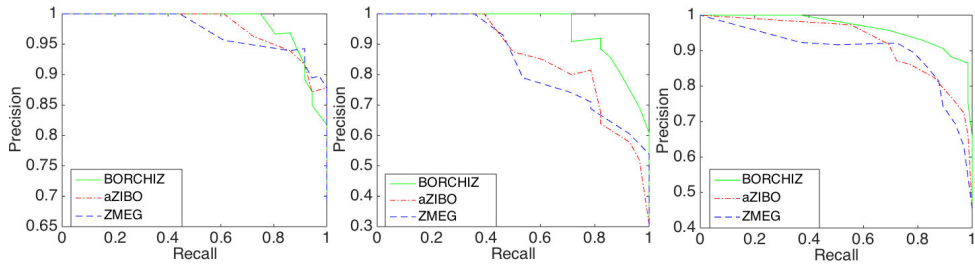


Figure 5.8: Precision-Recall curves (L-H classification) using B-ORCHIZ, aZIBO and ZMEG for Insert-C, Insert-I and the whole Insert dataset (from left to right).

Tool Wear Classification Based on the Wear Region Size

In this section, we evaluate the classification accuracy obtained using the wear size (area in mm^2). As in the previous experiment, we assess this approach on the Insert-C, Insert-I datasets and also on both together. In order to carry out a fair comparison with the results of our previous experiments, the classification stage was performed using a SVM with intersection kernel and following stratified 10-fold cross validation.

Figure 5.9 shows the classification accuracy for the L-H and L-M-H scenarios compared with those obtained with the B-ORCHIZ descriptor. In both experiments

B-ORCHIZ achieves the best performance. In the L-H classification, the best result was obtained in the Insert-I with an accuracy of 88.46% against 85.71% achieved with the simple area descriptor which involves a 3.21% of accuracy increment. In the three class classification setting the difference between these two approaches increases significantly. For the Insert-C dataset, B-ORCHIZ improvement reaches 23.35%. However, the highest difference appear when the experiments run on the full dataset (Inserts). The high difference of size between the complete edges and the incomplete ones makes the classification based on the area very poor with just a 61.63% of accuracy in the L-M-H dataset and less than 80% in L-H compared with the 80.24% and 87.06% achieved by our shape descriptor.

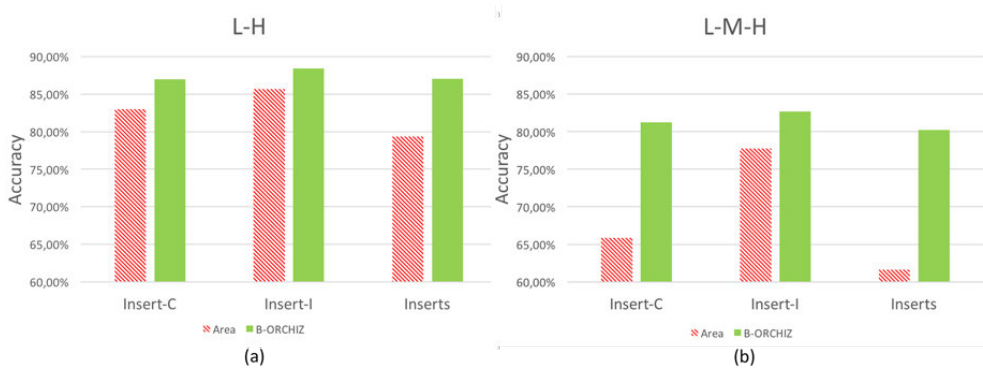


Figure 5.9: Classification results using B-ORCHIZ descriptor and the wear area size. In (a) the results for L-H experiment. In (b) the results for L-M-H experiment.

These results demonstrate the fact that the area information is not enough to describe with precision the wear level. Techniques based on shape description appear to be a better option to assess the wear level of the inserts.

Dendrogram analysis

A dendrogram is a tree-structured graph that represents the result of a hierarchical clustering as the distance between the different clusters.

In problems where the number of clusters is not fixed or when the goal is to discern between different sub-clusters within each class, the dendrogram may improve the capacity of taking correct decisions.

In addition, dendrograms give specific information about the correlation between the clusters: highly correlated groups are near the bottom of the graph, whilst as we move up the graph the clusters are bigger and the distance between them increases.

Automatic clustering for Insert dataset

We carried out an automatic hierarchical clustering to compare the groups obtained with the ground truth generated by the expert. The clustering was based on a hierarchical cluster tree using Ward linkage method (Ward, 1963). This analysis was conducted for each of the three evaluated descriptors, i.e. ZMEG, aZIBO and B-ORCHIZ and for the whole Insert dataset and the Insert-C and Insert-I subsets. The dendrograms which represent these clusterings are depicted in Figure 5.10.

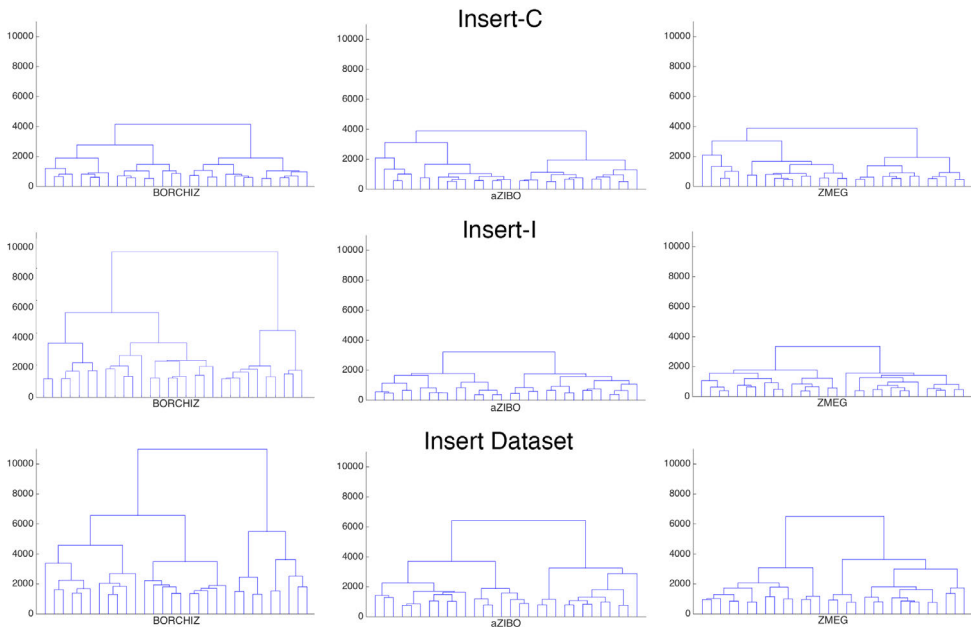


Figure 5.10: Dendrogram of the clustering obtained using B-ORCHIZ (left), aZIBO (center) and ZMEG(right) descriptors for the Insert dataset and the Insert-C and Insert-I subsets.

Our goal was to identify two and three classes within all the wear images.

A higher distance determines more separated clusters and therefore, a clearer distinction of whether an insert is damaged or not.

In the Insert-C dataset, the class division has a low distance between classes for all the descriptors. It is near 3200 for two classes and around 2000 for three classes. However, for the Insert-I subset, the distance between two classes is approximately 6000 with B-ORCHIZ, but just around 2000 with aZIBO and ZMEG. For three classes, the distance between the nearest clusters is around 1500 in both aZIBO and ZMEG and higher than 4000 for the B-ORCHIZ based clustering. This shows that B-ORCHIZ has more discriminant power in the Insert-I subset.

As it is shown in the third row of Figure 5.10, the study of the dendrogram on the combined set demonstrate that B-ORCHIZ generates descriptors with more discriminant power due to the higher difference between the classes. Ward linkage method ensures very homogeneous classes for the dataset and the dendrogram confirms this fact, showing a similar distribution of the classes.

According to this analysis, it can be stated that the best group separation is achieved using the proposed B-ORCHIZ method to describe the wear regions.

Representative shape of wear classes: Average greyscale image

After the automatic labelling process using ZMEG, aZIBO and B-ORCHIZ descriptors, we have carried out an evaluation of all the images that, according to the clustering process, belong to the same class in order to extract a representative shape for each one. The process is the following:

1. Firstly, all the edges have been aligned to the center of the image. For the Insert-C subset, this process consists simply on translating the region of the edge vertically so that its central row is in the middle row of the image. In the case of the Insert-I subset, the left and right edges have different orientations depending on the insert. Therefore, before translating them to the middle of the image we have to rotate all the incomplete regions to place them perfectly vertical.
2. Secondly, all the images have been resized to fixed dimensions and the element-wise sum for all the pixels of every image has been carried out. As we have binary images, if the resulting image has a pixel with value t , it means that t images have wear information on that position.
3. Finally, the resulting image is averaged and is shown as a grayscale image.

With that representative image for each cluster, we were trying to improve the comprehension of the wear edge classification and gave a clue to human experts about the estimated shape of each one.

In Figure 5.11 we can see the representative images for the three levels classification for the Insert-C subset. While aZIBO and ZMEG shows very similar representative images, the B-ORCHIZ high wear prototype is very different: B-ORCHIZ shows a larger wear area, what can be interpreted as the reason to make possible the best classification of that kind of images, allowing only images with high wear to be clustered together and leaving the rest for the medium or low wear cluster.

The same happens with the Insert-I subset where the difference between classes using B-ORCHIZ is higher than using the grouping provided by the expert.

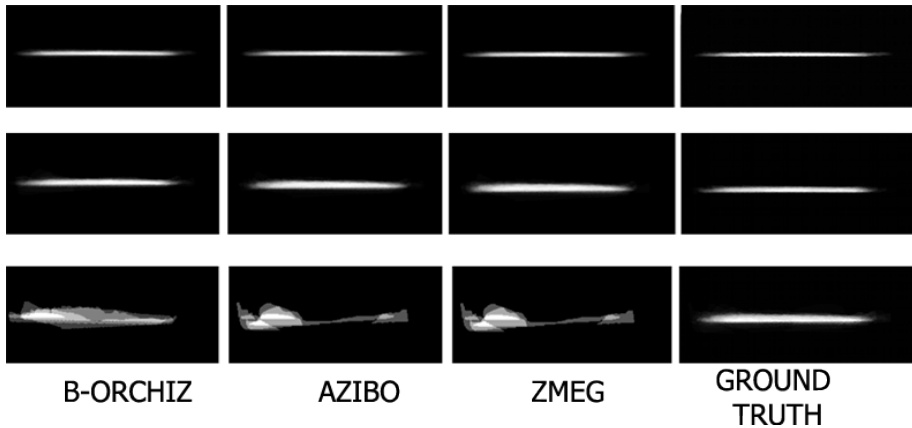


Figure 5.11: Representative images for three levels of wear (L: first row, M: second row and H: third row) for the Insert-C subset using clustering. Insert wear described with B-ORCHIZ, aZIBO and ZMEG with its associated ground truth (columns from left to right) for automatic clustering. Images are aligned and the average value of each pixel of the image representing the gray level of the representative image is calculated.

Based on this study, we can conclude that the descriptor that better represents the different edge wear is B-ORCHIZ because it provides more information about the wear region shape than the others descriptors and the expert clustering, being quite useful to study the wear process, to train new personal and to generate automatic templates.

Representative shape of wear classes: 3D Representation and contour map

In this section, we show with more detail the representative shape for each subset of the dataset (Insert-C and Insert-I) depending on the wear level. We focused on B-ORCHIZ to obtain the representative image for each class because, as we pointed out in the previous section, B-ORCHIZ offers more information about the shape depending on the wear level. Figure 5.12 depicts the representative images for the Insert-C and Insert-I when we consider two wear classes. In Figure 5.13, the representative images for the three classes are shown. These images demonstrate that the shape of the wear region can clearly offer reliable information about the insert state.

In general, the L-H labelling offers less information than the L-M-H one. Using the L-M-H labelling, L and M classes are quite similar in terms of shape but in the case of the H class, the shape is really irregular and the surface is bigger than in the other cases. It is also remarkable that, for the Insert-I dataset with low wear, the insert is more worn when it is closer to the cutting surface. Thus, the shape of the

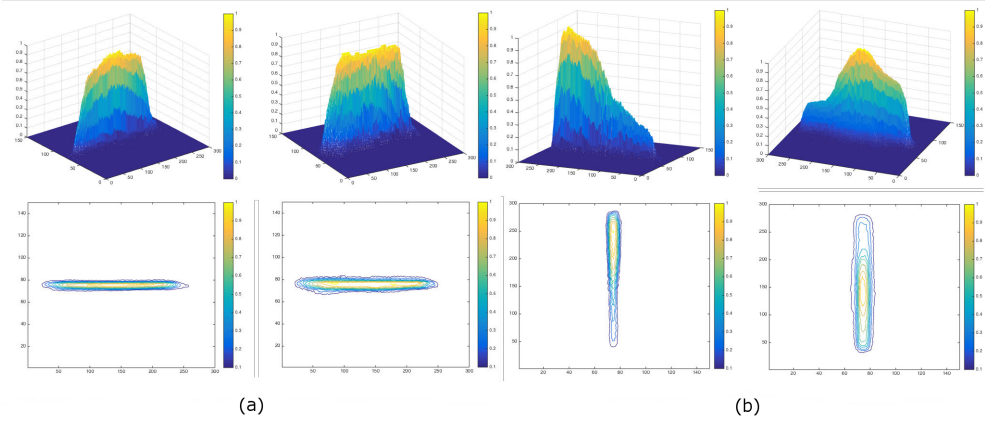


Figure 5.12: Representative images for (a) the Insert-C subset and (b) the Insert-I subset, L-H labelling (from left to right).

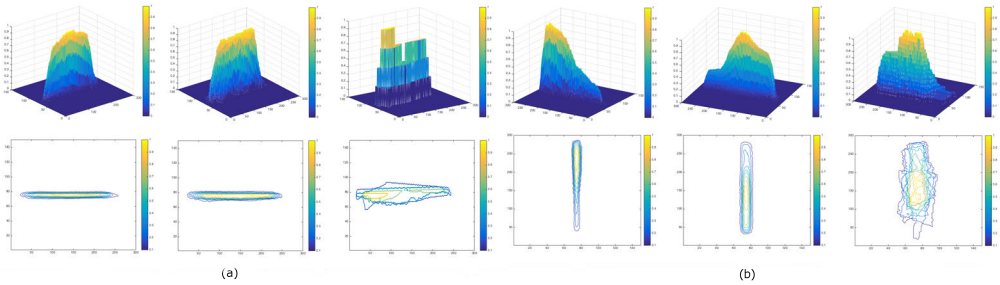


Figure 5.13: Representative images for (a) the Insert-C subset and (b) the Insert-I subset, L-M-H labelling (from left to right). First row represents the surface and second row represents the contour for the plan view.

wear region can clearly offer reliable information about the state of the insert.

5.4. Conclusions

In this chapter we proposed two new shape descriptors based on the ZMEG algorithm: aZIBO and B-ORCHIZ. Several experiments using well known datasets like Kimia or MPEG7 have been carried out comparing the performance of these three methods. Results shown that B-ORCHIZ improves the performance achieved by the original ZMEG method and also the aZIBO proposal. One of the main reasons of this improvement can be the total invariance to rotation of our methods

in contrast with the ZMEG algorithm which just is invariant to rotation in its global feature but not in the local ones.

Furthermore, we presented a new approach to monitor tool wear in edge profile milling operations using computer vision. Our aim was to classify the insert according to its wear stage as a function of the wear region shape with a machine learning classification model. To achieve this goal, we proposed B-ORCHIZ, a new shape descriptor computed from the tool image that captures both, global and local image information, and guarantees total invariance to rotation.

In order to assess this proposal, we created a dataset with images from 53 tools (with four cutting edges each one) that was labelled by an expert into two and three wear classes.

Following a supervised learning approach, the performance of a SVM classifier in terms of classification accuracy was assessed on the Insert dataset comparing B-ORCHIZ with other two descriptors: ZMEG and aZIBO. B-ORCHIZ outperformed the other methods in the scenarios evaluated. Additionally, a cluster analysis shown that B-ORCHIZ has more discrimination power to separate the tool wear images into homogeneous groups and it provides more detailed information of the wear region states.

These results shown that computer vision techniques can be taken into account for future real projects as a reliable approach that enables tool condition monitoring in edge profile milling operations.

Chapter 6

Fusion of region and moment descriptor

In this chapter, a new system based on the combination of a contour descriptor and a shape descriptor has been proposed for classifying inserts in milling processes. To describe the contour we have used the B-ORCHIZ method detailed in the previous chapter, because it has demonstrated well performance in several experiments carried out, achieving the best results of the state of the art in tool wear monitoring (García-Ordás et al., 2016). For the shape description we have proposed a new descriptor called ShapeFeat.

Our motivation to present ShapeFeat is that it takes into account the properties of the binary region instead of just contour information to describe it. Furthermore, it is just formed by ten features, which can be very appropriated for embedded systems or applications where the storage is a critical requirement. The fusion of B-ORCHIZ with this proposal was carried out using three different techniques increasing the discriminative power of both methods individually. Due to the fact that each descriptor uses different features of the image, a combination of both discriminates better all the possible differences between shapes.

6.1. Methodology

6.1.1. Shape descriptor: ShapeFeat

We referred to the proposed shape descriptor as ShapeFeat and it relies on ten different features extracted from the binarized image. This descriptor provides useful information about the object which is not usually obtained with moment descriptors such as Hu, Flusser, Zernike, etc. that are focused in the shape information, but avoiding other data such as the area of the region or the homogeneity of it.

The ten different features for describing the binary regions that we have selected are: ConvexArea, Eccentricity, Perimeter, EquivDiameter, Extent, FilledArea, FilledImage, MinorAxisLength, MajorAxisLength and Solidity.

- ConvexArea: Number of pixels of the smallest convex polygon that contains the region. Once we have the coordinates of the smallest convex polygon $((x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n))$, they are arranged in the determinant

below. The coordinates must be taken in counterclockwise order around the polygon, beginning and ending at the same point.

$$Area = \frac{1}{2} \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \dots & \dots \\ x_n & y_n \\ x_1 & y_1 \end{vmatrix} = \frac{1}{2} [(x_1y_2 + x_2y_3 + x_3y_4 + \dots + x_ny_1) - (y_1x_2 + y_2x_3 + y_3x_4 + \dots + y_nx_1)] \quad (6.1)$$

- **Eccentricity:** Scalar that specifies the eccentricity of the ellipse that has the same second central moments as the region (Figure 6.1). The eccentricity is the ratio of the distance between the foci of the ellipse and its major axis length. The value is between 0 and 1. (0 and 1 are degenerate cases: an ellipse whose eccentricity is 0 is actually a circle, while an ellipse whose eccentricity is 1 is a line segment.)

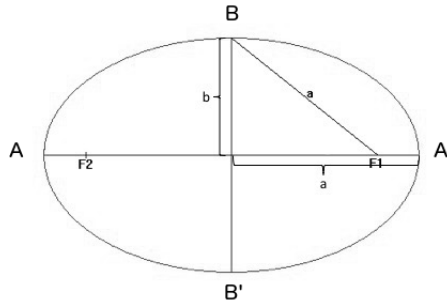


Figure 6.1: Ellipse with centre O . The foci coordinates are $F'(-c, 0)$ and $F(0, c)$

For all points in the ellipse: $\overline{PF} + \overline{PF'} = 2a$.

This expression results in:

$$\sqrt{(x - c)^2 + y^2} + \sqrt{(x + c)^2 + y^2} = 2a \quad (6.2)$$

Where we reach:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (6.3)$$

Then, the eccentricity of the ellipse is calculated as:

$$\sqrt{1 - \frac{b^2}{a^2}} \quad (6.4)$$

- Perimeter: Number of the boundary pixels.
- Equivalent diameter: Scalar that specifies the diameter of a circle with the same area as the region. It is computed as $\sqrt{\frac{4A}{\pi}}$, where A is the area.
- Extent: Scalar that specifies the ratio of the pixels in the region to the pixels in the total bounding box.
- FilledArea: Number of pixels belonging to the image after filling holes.
- MinorAxisLength: Length of the segment $\bar{B}B'$ shown in Figure 6.1 (2b).
- MajorAxisLength: Length of the segment $\bar{A}A'$ shown in Figure 6.1 (2a).
- $R = \frac{2b}{2a}$, where $2b$ and $2a$ stands for the lengths of the minor and the major axis, respectively, of the ellipse that has the same normalized second central moments as the region.
- Solidity: Scalar indicating the proportion of the pixels in the smallest convex polygon that can contain the region, that are also in the region. So it is computed as $\frac{Area}{ConvexArea}$.

In Figure 6.2 we can see some of the relevant information used in the proposed ShapeFeat descriptor. In (a), we can see the semi-minor and semi-major axes of the ellipse that has the same second central moments as the region. This information is also used to extract the eccentricity and R . In (b) we can see the boundary box image crop, used to obtain the Extent property just calculating the ratio between the number of pixels in that image and the number of white pixels (region pixels). In (c), the smallest convex polygon is shown. That polygon is necessary to extract the Convex area value and the Solidity, which is the ratio between the number of white pixels in (b) and the number of white pixels in (c).

6.2. Combination of Shape and Contour descriptors

Humans usually take into account different combination of features of an image to interpret them or to identify people, animals or objects. For example, yellow color can be useful to identify a banana in a image, but without consider shape properties, it can be mistaken with a lemon or the sun. This basic idea can be used in computer vision to increase the chance of success on image classification problems. In our

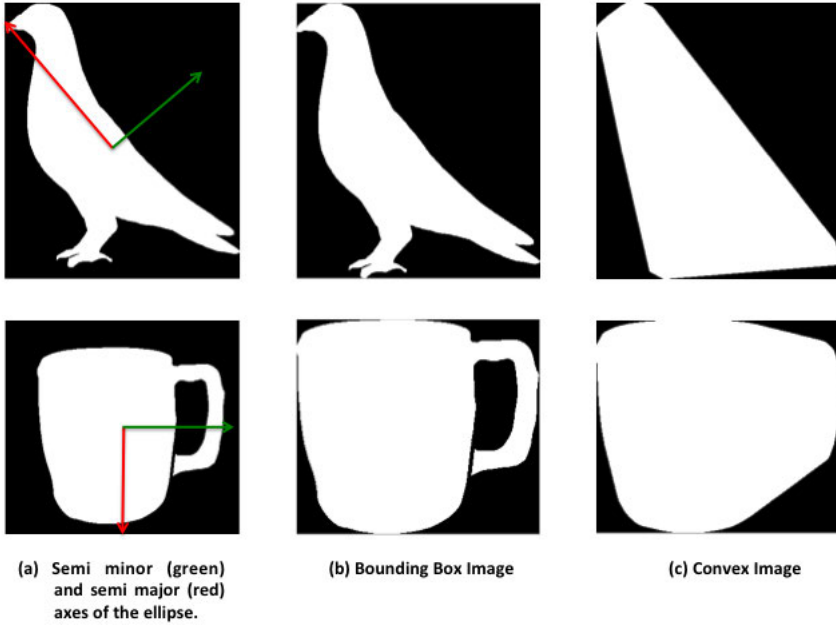


Figure 6.2: Two examples of the ShapeFeat features extraction. With the information shown in the first column, we can extract the eccentricity and R. With the information shown in the second one, we can extract the extent and finally, the smallest convex polygon shown in the third column, is necessary to extract the Convex Area and the Solidity.

work, we explore the combination of shape and contour features to improve our tool wear monitoring system. There are three well-known ways to merge descriptors: early fusion, intermediate fusion and late fusion. (See Figure 6.3). Next, we describe with more detail each of them.

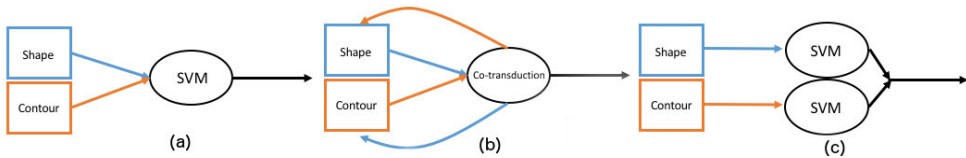


Figure 6.3: (a) Early fusion of the contour and shape descriptors. (b) Intermediate fusion using co-transduction method. (c) Late fusion using Bayes average.

6.2.1. Early fusion

Early fusion combines the features of different techniques and produces a single feature to the classifier. The process inherently increases the size of feature vector but it is one of the fastest ways of fusion. In our case, we have used the simple concatenation of feature vectors. Thus, our input descriptor will be composed of B-ORCHIZ and ShapeFeat features.

6.2.2. Intermediate fusion: Co-transduction method

Co-transduction method is a retrieval system developed by Xiang bai et al Bai et al. (2012). Their goal was to develop an algorithm to fuse different similarity measures for robust shape retrieval through a semisupervised learning framework. The method was named Co-transduction, which is inspired by the Co-training algorithm (Blum and Mitchell, 1998). In our case, the method has been adapted to deal with classification instead of retrieval using a k-Nearest neighbors approach with the top k retrieved images. Furthermore, we have combined two different descriptors instead of two different similarity measures to take advantages of all the information extracted by ShapeFeat and B-ORCHIZ. Given a query shape the algorithm iteratively retrieves the most similar shapes using one description similarity matrix and assigns them to a pool for the other one to do a re-ranking, and vice versa.

The following is the pseudocode of Co-transduction method:

Input: a query object x_1 , the collection of objects $X = x_2, \dots, x_n$.

Process:

Create a $n \times n$ probabilistic transition matrix P_1 based on one type of shape similarity (BORCHIZ in this case).

Create a $n \times n$ probabilistic transition matrix P_2 based on one type of shape similarity (ShapeFeat in this case).

Create two sets Y_1, Y_2 such that $Y_1 = Y_2 = x_1$

Create two sets X_1, X_2 such that $X_1 = X_2 = X$

Loop for m iterations:

Use P_1 to learn a new similarity sim_1^j by graph transduction when Y_1 is used as the query objects ($j = 1, \dots, m$ is the iteration index)

Use P_2 to learn a new similarity sim_2^j by graph transduction when Y_2 is used as the query objects

Add the p nearest neighbors from X_1 to Y_1 based on the similarity sim_1^j to Y_2

Add the p nearest neighbors from X_2 to Y_2 based on the similarity sim_2^j to Y_1

$X_1 = X_1 - Y_1$

$X_2 = X_2 - Y_2$

(Then, X_1 and X_2 will be unlabeled data for graph transduction in the next iteration).

The method is widely explained in Bai et al. (2012).

6.2.3. Late fusion

The late fusion approaches use multiple classifiers to determine the output instead of just one as early fusion does. These methods try to combine all the prediction scores, which indicates the confidence of classifying the sample as positive. Although very simple, this method has proved to be effective in improving performance of each individual classifier.

In this dissertation we have used the simple Bayes average (Ruta and Gabrys, 2000; Bostrom, 2007) as a method for obtaining a class probability distribution from the fused classifiers trained using contour (B-ORCHIZ) and shape (ShapeFeat) descriptors:

$$P_{LATEFUSION}(x \in C|x) = \frac{P_{SHAPEFEAT}(x \in C|x) + P_{BORCHIZ}(x \in C|x)}{2}, \quad (6.5)$$

where x is one image and C one of all the possible classes it may belong to. $P_{SHAPEFEAT}$ and $P_{BORCHIZ}$ are the probability distributions for the different classifications carried out.

This method takes into account the information provided by all the classifications and determine a decision based on the mean average of all of them, which avoid problems derived from a possible not fair classification.

6.3. Tool Wear monitoring

6.3.1. Experimental setup

To determine the wear of the complete edge by itself and also how it is influenced by the incomplete edges, we divided the whole Insert dataset into two: (a) Insert-C subset with the complete edges that correspond with the horizontal wear in the cropped images, and (b) Insert-I subset with the incomplete ones which orientations are vertical in the images.

To carry out the classification for these three datasets with all the descriptors evaluated, we used a SVM classifier with intersection kernel. We have used Monte Carlo cross-validation, which randomly splits the dataset into training and test data. For each split, the model is fit to the training data, and predictive accuracy is assessed using the test subset. The final result is the average of all runs (20 in our case). The advantage of this method (over k-fold cross validation) is that the

proportion of the training/test split is not dependent on the number of iterations. In our work we have used a 75%-25% of train - test subsets. Results are shown in terms of the accuracy achieved in the classification.

We combined the two best descriptors (B-ORCHIZ and ShapeFeat) using early, intermediate (co-transduction) and late fusion techniques and results were compared with each of them separately and also with other classical descriptors. Co-transduction classifier uses k-NN with k equals 3, 7, 9 and 11 obtaining the best results for k=3 in all cases.

6.3.2. Insert Region dataset results

On this section, we are going to show the results obtained in all of our experiments divided in two main categories. Results achieved by each descriptor by itself and results combining both of them using different fusion techniques: early fusion, intermediate fusion and late fusion.

Results without fusion

The first step consists of assessing the descriptor based on shape presented in this paper (ShapeFeat) and other proposals based on contour: B-ORCHIZ, ZMEG and aZIBO.

Table 6.1, shows the performance in terms of accuracy

Table 6.1: Classification accuracy in % of ZMEG, aZIBO, B-ORCHIZ and ShapeFeat using SVM with Intersection kernel for the complete, Insert-C, Insert-I dataset (from left to right) for two and three wear levels.

	Complete		Insert-C		Insert-I	
	L-H	L-M-H	L-H	L-M-H	L-H	L-M-H
ZMEG	83.74	75.87	85.58	76.44	83.52	79.40
aZIBO	84.44	78.85	87.02	76.92	84.89	82.14
B-ORCHIZ	87.06	80.24	87.02	81.25	88.46	82.69
ShapeFeat	88.70	80.67	93.37	81.35	88.41	84.12

It can be seen that B-ORCHIZ achieves in all the datasets and all the classifications better results than ZMEG and aZIBO descriptors. The improvement for the complete dataset was of more than a 3% in the binary classification and almost a 2% in the ternary one with respect the best descriptor of the state of the art (aZIBO). However, these results were outperformed with our shape proposal (ShapeFeat) in almost all the experiments. As we can see, the highest improvement was achieved with the Insert-C dataset, with an increment of more than a 7.29% of

accuracy. In all the other cases, the behavior was very similar in comparison with the contour methods.

Results with fusion of contour and shape descriptors

B-ORCHIZ (contour) and our ShapeFeat (shape) are, for the moment, the methods obtaining best results in the scenario of using digital image for assessing tool wear. The high difference of both methods in the way they are built, allows us to explore about different fusion techniques combining both of them. In this section, we show the results of three different fusion techniques applied over these two methods: early, intermediate and late fusion.

Early fusion was performed by concatenating B-ORCHIZ and ShapeFeat to create a new feature vector used as input for the SVM classifier. Intermediate fusion was carried out by using the previously explained classifier called Co-Transduction. Late fusion was implemented combining the scores of the SVM classifier for B-ORCHIZ and ShapeFeat in order to determine the final response of the learner.

In Figure 6.4, results for each fusion method are shown. In order to compare the results with the original methods, we have also included B-ORCHIZ and ShapeFeat. As we can see, early fusion method achieves a very similar result as the B-ORCHIZ method by itself. This can be explained due to the high difference of size between both descriptors, which makes B-ORCHIZ predominate over ShapeFeat. Co-Transduction and Late fusion are invariant to the number of features because the fusion is carried out after the classification step. In almost all the experiments late fusion demonstrates a higher performance than the rest of the description techniques. However, in the binary classification using the Insert-C dataset, the high difference of accuracy between shape and contour descriptors conditions the result of the early and late fusion methods. The good performance of the co-transduction algorithm for this experiment is obtained because of the fusion algorithm. Whereas late fusion and early fusion are averagely influenced by both performances, in co-transduction algorithm each descriptor is improved by the other. For this reason, although B-ORCHIZ do not show good results, it also improves the performance of ShapeFeat instead of decrease it in the fusion step.

Descriptor fusion vs baseline descriptors

We also compare our proposal with other classical descriptors like BCF (Bag of contour fragments), HOG (Histogram of oriented gradients) and SC (Shape Context) (Wang, Feng, Bai, Liu and Latecki, 2014; Dalal and Triggs, 2005; Belongie et al., 2002). In Table 6.2, the superiority of our combined method is underlined outperforming the state of the art methods in all the cases, achieving improvements of more than 82% respect SC, more than 35% respect HOG and 31% of improvement taking

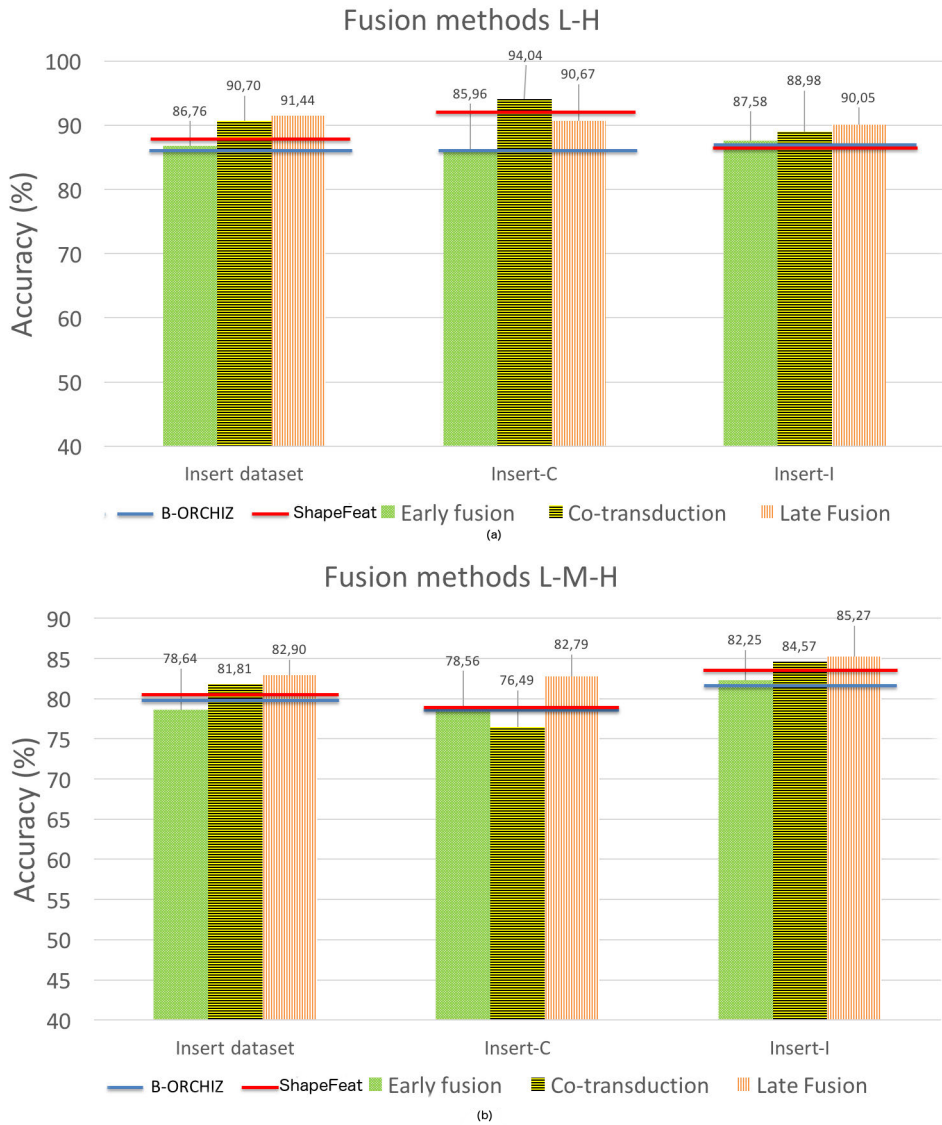


Figure 6.4: Results for binary (a) and ternary (b) classification using fusion methodology with B-ORCHIZ and ShapeFeat. Furthermore, results achieved by the descriptors by themselves are represented as straight lines over the graphs.

into account the BCF method, in all the cases evaluating the L-M-H classification performance for the complete dataset.

Table 6.2: Classification accuracy in % of the combination of ShapeFeat with BORCHIZ and other descriptors like BCF (Bag of contour fragments), Hog (Histogram of oriented gradients) and SC (Shape Context)

	Complete		Insert-C		Insert-I	
	L-H	L-M-H	L-H	L-M-H	L-H	L-M-H
Late fusion	91.44	82.90	90.67	82.79	90.05	85.27
BCF	76.76	63.11	77.88	49.52	80.77	74.45
HOG	76.80	60.35	80.19	51.06	76.65	66.15
SC	54.58	45.45	68.75	41.83	69.51	48.08

6.4. Conclusions

In this chapter, a new method for automatic evaluation of the wear level in milling inserts have been developed. Evaluating the most recent studies connected with this field and the shape feature extraction, up to the moment, the best technique in the state of the art for evaluating the status of the inserts is the B-ORCHIZ method. B-ORCHIZ method is based in the contour description of the shape. In order to evaluate other possibilities, we proposed to use ten shape features properties of the binary images.

Taking into account the good performance of the contour and shape descriptors, we proposed to combine both methods. The fusion was carried out using three different fusion methods: early, intermediate (which combines the similarity matrices of each method to improve the classifier of the other one) and late fusion. The results achieved were 91.44% of hit rate for binary classification using the whole dataset and 82.90% in the low-medium-high one combining both methods with late fusion. These results show an opportunity for wear monitoring automation in edge profile milling processes, saving time in the process of analyzing the insert, its associated costs and the possible errors due to the subjectivity of the humans evaluation.

Chapter 7

Patches based descriptor

As we stated in the two previous chapters, in the context of tool replacement operations, indirect costs derived from the unproductive time needed to perform the tool replacement are significant. A portable system to evaluate the inserts is of great interest in the context of tool replacement operations. For this purpose, we proposed a descriptor based on patches.

Some works have used description of different patches extracted from an image for different object recognition tasks such as 3D face recognition (Li et al., 2015) or pedestrian detection (Halidou et al., 2014; Hoang et al., 2014) merging the features extracted from each block. Instead of merging different features coming from the same patch, here we describe and classify each patch individually, so the system is able to provide an estimation of the degree of tool wear by aggregating these individual classifications. Note that this gives robustness to the final decision, as a possible misclassification of a few single blocks would not affect the final tool wear estimation system as much as if the method misclassifies the whole image. Furthermore, using such multi-block approach, our system is able to detect small regions of high wear in the image which might have been ignored otherwise. It is important to highlight that our system does not require any segmentation of the tool wear region, which is a very time consuming step and, moreover, might not always be accurate enough.

This method can be integrated in a system and we refer to this process as portable, meaning that it is a system whose dimensions are small and its cost is low. Such systems, built upon Single Board Computers (SBC) (e.g. Raspberry Pi or Odroid), make possible to implement solutions easily in embedded systems with low cost and low power consumption, and time performances comparable to those obtained using traditional computers.

It is also important to highlight that in a real environment the tool wear assessment is carried out while the head tools are resting, which takes between 5 to 30 minutes (Fernández-Robles et al., 2015b). The proposed system takes approximately 0.13 seconds to evaluate each insert, so it takes around 4 seconds to assess the entire head tool (including the time to rotate it), which have 30 inserts.

Some works that apply computer vision and portable systems have been proposed recently. For example, Tu et al. introduced a system for honeybee counting

using a Raspberry Pi and computer vision techniques (Tu et al., 2016). Other applications have also been developed recently with embedded systems and low cost cameras (Shakshuki et al., 2015; Cambuim et al., 2016) but, to the best of our knowledge, this is the first time that a portable system is proposed to tool wear monitoring.

Nowadays, if we focus in digital image processing, texture methods are one of the most used due to their high discriminative power and low computational cost (Ojala and Pietikainen, 2002). For this reason we proposed a new method based on local texture description trying to improve the previous contour, moment, orientation and region descriptors.

The aforementioned method is based on image texture analysis for tool wear monitoring in an edge profile milling machine and it can be embedded in a portable system. Our approach is based on obtaining local texture features individually from several regions extracted from the zone of the tool where the wear tends to appear. This approach presents two main advantages: firstly, establishing wear patches (WP) with different sizes and orientations allows to detect small - but important- worn areas that otherwise (i.e. using methods that extract a single feature vector from the whole image) would have been overlooked. Secondly, it avoids the segmentation stage, which saves time and computational resources, making feasible a low cost portable implementation. Additionally, since each wear patch classification is addressed individually using supervised learning techniques based on kernels (i.e., SVM), the monitoring system is able to provide an estimation of the tool wear percentage by aggregating the individual results.

The highest improvement of this system with respect to the previous methods evaluated is not just the high classification rates but also the capability of being full automatic and a very low cost system.

7.1. Methodology

The outline of the whole methodology proposed in this chapter for tool wear monitoring in edge profile milling processes is shown in Figure 7.1 where the training and operational stages are shown. Next, we briefly describe the whole process and in the following sections, we elaborate on each one.

First of all, a portable system is used to take gray scale images of the tool head. Afterwards, the cutting edges are automatically detected and extracted. Then, each cutting edge is divided in several wear patches (WP) and each one is characterised using texture descriptors based on Local Binary Pattern (LBP). Finally, each WP is classified using a Support Vector Machine (SVM), and these classifications are used to make a final decision about the wear state of the tool cutting edge.

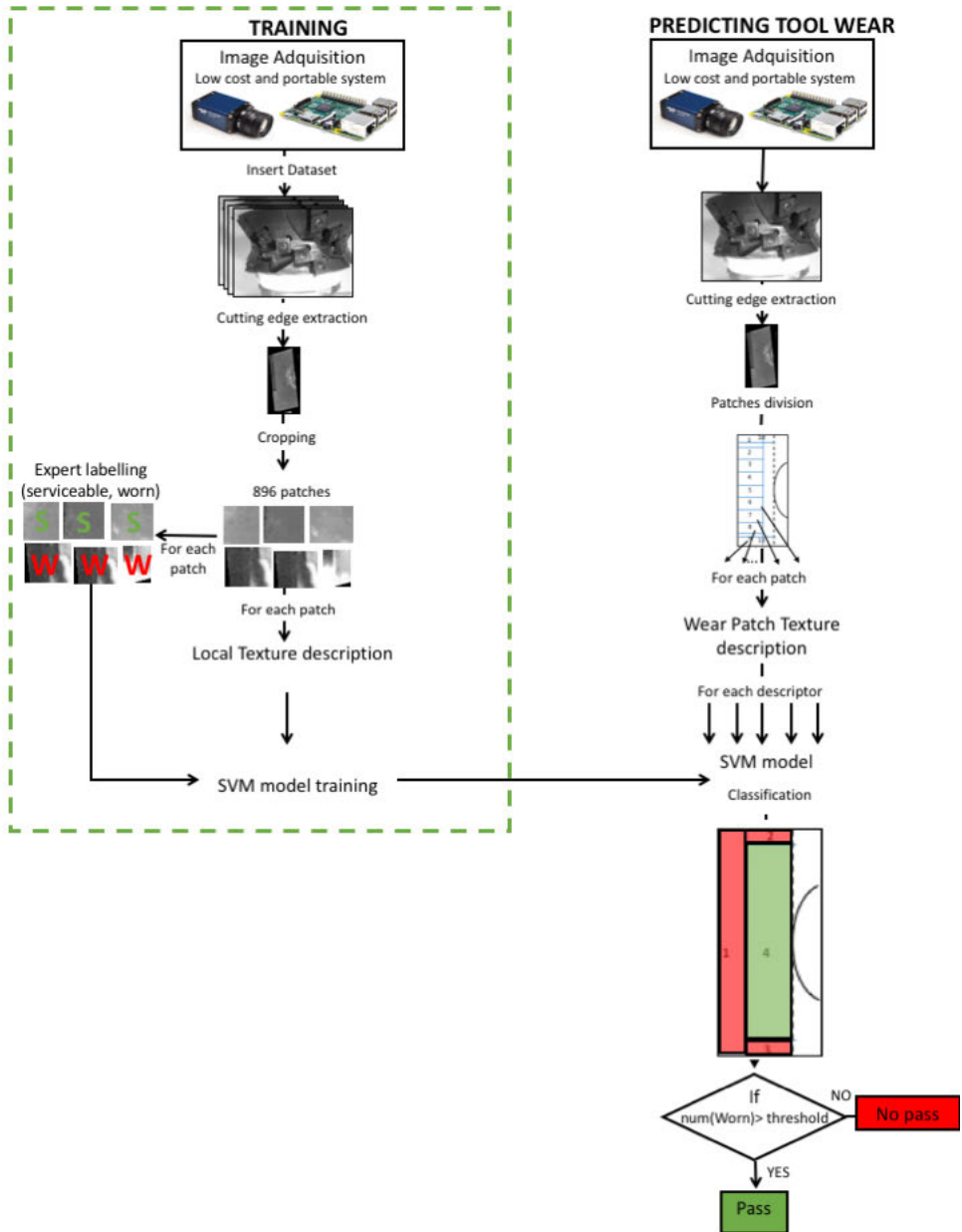


Figure 7.1: Schema of our proposed system for tool wear monitoring using computer vision.

7.1.1. Regions configuration

In this section, we present different alternative divisions trying to adapt the patches' shape and position to our region of interest. A graphical representation is shown in Figure 7.2 and next, we describe each one with more detail. It is a knowledge driven patch configuration (i.e. the cutting edge image division) considering both the typical position and shape of the WP.

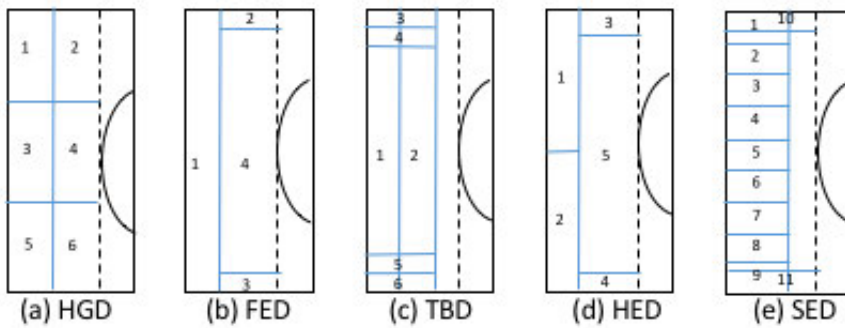


Figure 7.2: All the different configurations of the patches over the cutting edges.

Homogeneous Grid Division (HGD)

In this case the cutting edge image is divided into a 3×2 grid. With this configuration, we create patches with the same size in order to give the same importance to all of them in the classification step. The Homogeneous Grid Division (HGD) is the only configuration that has uniform patch size (See Figure 7.2(a)).

Full Edge Division (FED)

The Full Edge Division (FED) is based on the basic idea of evaluating each cutting edge of the image individually. Figure 7.2(b) shows a representation of this configuration. For each image, four different patches have been extracted: The main cutting edge patch, two horizontal patches connected with the upper and lower part of the insert, which usually have also some kind of erosion, and finally the interior region of the insert. In this proposal, all the patches are fully independent and includes all the information which appears in the cutting edge image.

Two Bands Division (TBD)

The TBD alternative, shown in Figure 7.2(c), extracts six patches per image. The first one, as we did in FED, contains the information of the main cutting edge.

Additionally, a patch of the same size as the first one located after the first patch is taken into account in order to capture possible high wear regions located also after the border of the first patch division. We proceed the same way with the horizontal patches, extracting two different patches in the upper zone and two more in the lower one. In this case, the horizontal and vertical patches are overlapped.

Half Edge Division (HED)

The Half Edge Division (HED) proposal (see Figure 7.2(d)), has the same patches configuration as FED, but dividing the vertical patch that contains the edge information into two: The first one containing the information of the main cutting edge located in the top of the insert and the second one with the information of the bottom. With this division, we detect small wear regions that could be misclassified in the first approach due to their specific location in just one part of the vertical patch.

Small Edge Division (SED)

Trying to merge some of the ideas used in the other configurations, we propose the Small Edge Division (SED). In the same way as in the TBD method, we want to extract the information of the high wear region but, instead of using two vertical bands, we use a wider region. Furthermore, we have taken into account the idea of the HED method, which divides the vertical patch in two different regions but in this case we have gone further dividing the vertical region in 9 subregions, as we can see in Figure 7.2(e). Moreover, we have also evaluated the top and bottom parts of the insert as we did in FED, but starting from the corner of the main edge which causes an overlapping between these patches and the two extreme sub-regions of the vertical one.

7.1.2. Texture descriptors

Local Binary Pattern (LBP)

Local Binary Pattern (LBP) (Ojala and Pietikainen, 1996) is a gray-scale texture operator that extracts pixel-wise information of an image. For each pixel c , LBP takes into account its P neighbours within a radius (i.e. distance) R . When a neighbour p has a grey level (i.e. intensity) value g_p greater than or equal to that of c , the value 1 is assigned to it, or 0 otherwise.

Thereafter, the LBP for that pixel is calculated by summing up those values multiplied by consecutive powers of 2, as it is stated in Equation (7.1). Figure 7.3 depicts an example of the extraction of the LPB of one pixel.

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c) 2^p, \quad s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}, \quad (7.1)$$

where g_c is the intensity value of central pixel, P is the number of neighbours, g_p is the value of its p -th neighbour, which lies along the orientation $2\pi p/P$ at a distance R (i.e. the radius of the neighbourhood).

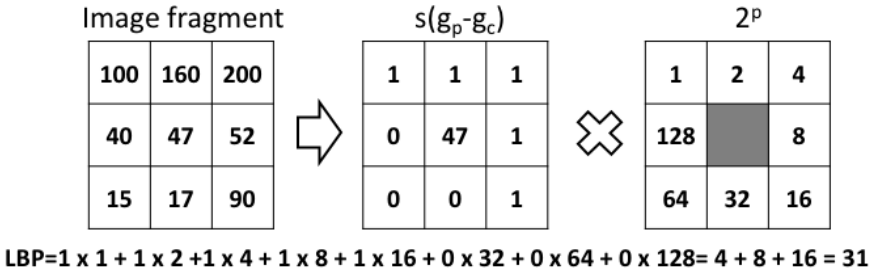


Figure 7.3: Local Binary Pattern process over one pixel in gray scale level using a neighbourhood of radius 1 and 8 neighbours.

Despite LBP is invariant to monotonic transformations of the gray scale, rotations may yield different LBP values. Therefore, Ojala et al. (Ojala and Pietikainen, 2002) defined a new formulation of LBP to achieve invariance to rotation by assigning a unique identifier to each rotated LBP. This new definition is stated in Equation (7.2).

$$LBP_{P,R}^i = \min \{ ROR(LBP_{P,R}, i) \mid i = 0, 1, \dots, P-1 \}, \quad (7.2)$$

where $ROR(x, i)$ is an operator that performs a circular bit-wise right shift i times on the P -bit number x . For example the rotation invariant patterns of 11001011 and 10010111 have the same value, i.e. 00101111.

Adaptive Local Binary Pattern (ALBP)

Adaptive Local Binary Pattern (ALBP) is a descriptor based in LBP proposed by Guo et al. (Guo, Zhang, Zhang and Zhang, 2010). It is motivated by the lack of information about the orientation of conventional LBP. The oriented mean and standard deviations of the local absolute differences $|g_c - g_p|$, $\forall p \in \{0, 1, \dots, P-1\}$ are taken into account in order to make the matching more robust against local spatial structure changes.

Intuitively, the texture classification could be improved by minimising the variations of the mean and standard deviation of the directional differences.

Guo et al. introduced a parameter w_p so that the p -th directional difference $|g_c - w_p \cdot g_p|$ can be minimised, and defined the following objective function:

$$w_p = \underset{w}{\operatorname{argmin}} \left\{ \sum_{i=1}^N \sum_{j=1}^M |g_c(i, j) - w \cdot g_p(i, j)|^2 \right\}, \quad (7.3)$$

where w_p is the weight used to minimise the directional difference, and N and M are the number of rows and columns in the image, respectively. Guo et al. used the least squares estimation method to minimise such objective function and, thus, to obtain the optimum weight parameter vector, i.e. $\bar{w} = [w_0, w_1, \dots, w_{P-1}]$.

Finally, ALBP is defined as follows:

$$ALBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - w_p \cdot g_c) 2^p, \quad s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}. \quad (7.4)$$

Completed Local Binary Pattern (CLBP)

Completed Local Binary Pattern (CLBP) was proposed by Guo et al. (Guo, Zhang and Zhang, 2010a) in order to generalize and complete the classical LBP. In this case, given a central pixel c with intensity g_c and its P neighbours with intensities g_p the local difference vector (i.e. $[d_0, d_1, \dots, d_{P-1}]$, where $d_p = g_p - g_c$, $\forall p \in 0, 1, \dots, P-1$), which characterises the image local structure at g_c , is taken into account. This local structure is then represented by means of the Local Difference Sign-Magnitude Transform (LDSMT), which decomposes each d_p into two components: its sign s_p (i.e. 1 if $d_p \geq 0$ or -1 otherwise), and the magnitude m_p (i.e. $|d_p|$).

Two operators are proposed to model the signs (S) and the magnitudes (M) of the local differences, namely CLBP_S and CLBP_M. The CLBP_S operator is calculated in the same way as LBP histogram (see Equation (7.1)). CLBP_M is defined in Equation (7.5).

$$CLBP_M_{P,R} = \sum_{p=0}^{P-1} t(m_p, a) 2^p, \quad t(x, a) = \begin{cases} 1 & \text{if } x \geq a \\ 0 & \text{if } x < a \end{cases} \quad (7.5)$$

where a is a threshold determined adaptively. In this case we have set it as the mean value of m_p .

Finally, since both operators are in binary format, they can be concatenated to form the final CLBP histogram.

Local Binary Pattern Variance (LBPV)

LBP Variance (LBPV) (Guo, Zhang and Zhang, 2010b), is another proposal made by Guo et al. which combines LBP and a contrast distribution. First, the uniform LBP (Ojala and Pietikainen, 1996) of the image is calculated. Then, the local variances of the image are used as a weight to adjust the contribution of the LBP code in the histogram calculation. The LBPV histogram is computed as:

$$LBPV_{P,R}(k) = \sum_{i=1}^N \sum_{j=1}^M w(LBP_{P,R}(i,j), k), \quad k \in [0, K], \quad (7.6)$$

where k is each bin of the histogram, K the maximum value of LBP and w is defined as:

$$w(LBP_{P,R}(i,j), k) = \begin{cases} VAR_{P,R}(i,j), & LBP_{P,R}(i,j) = k \\ 0 & \text{otherwise} \end{cases} \quad (7.7)$$

In Equation (7.7) $VAR_{P,R}$ is the rotation invariant measure of the local variance (i.e. the variance of a neighbourhood), defined as:

$$VAR_{P,R} = \frac{1}{P} \sum_{p=0}^{P-1} (g_p - u)^2, \quad (7.8)$$

where $u = 1/P \sum_{p=0}^{P-1} g_p$.

7.2. Tool Wear monitoring

7.2.1. Experimental setup

The goal of this section is to create a model that allows to assess individual patches and, based on its predictions, makes a final decision about the tool wear level.

Our high resolution dataset is composed by 577 images of the cutting edge. In our experiments, the set of cutting edge images has been randomly divided into a training set and a test, which contains 70% and 30% of the whole set, respectively.

Each image of the training set was divided into patches that later on were labelled by an expert in the field as belonging to a worn or functional region. This has been done manually. The manual division avoids possible bad extractions of the patches that could lead to the generation of suboptimal classifiers.

After this process we have obtained 896 patches, being 466 of them serviceable and 430 worn. We have used the descriptors of the patches of the training subset to

model a Support Vector Machine (SVM) classifier that uses an intersection kernel.

Finally, we automatically extract the patches for each image of the test subset using the methods explained in Section 7.1.1. All these patches were described using the methods based on LBP explained previously. After that, each patch was classified using the SVM model in order to estimate their labels (i.e. functional or worn) and, finally, we calculated the proportion of deteriorated patches in the analysed insert.

We established a threshold parameter in order to determine if an insert is still disposable: if an image contains a number of patches classified as worn higher than the threshold, then the tool edge is considered to be worn. The threshold can vary between one and the number of patches into which the image is divided. The higher the threshold, the stricter it is for the method to label an insert as disposable.

7.2.2. Threshold selection

First of all, we have evaluated all the descriptors with all the different wear region configurations varying the number of wear patches (WP) that are necessary to consider an insert as worn. In these experiments, we have focused in the recall metric because of the nature of our problem: the impact of misclassifying a worn cutting edge image as serviceable is higher than the other way round.

Figure 7.4 shows the recall values for all the different experiments using all the assessed threshold values. As expected, in all the cases, the threshold equal to one offers the best recall results with any descriptor and wear region configuration. The higher the threshold value is, the lower the recall achieved with our approach. However, not all the wear region configurations show the same behaviour: the performance using FED, TBD and HED is very dependent on the threshold. For example, using $ALBP_{16,2}$ and FED, the recall varies from 0.8 when the threshold is one to less than 0.2 with threshold equal to four. In contrast, HGD and SED achieve more invariance to the threshold with certain descriptors. For example, the difference with the same thresholds and descriptors is approximately 0.1 in HGD with $ALBP_{16,2}$ on the same interval. This may be due to the higher number of patches that are extracted in these two wear region configurations: the higher the number of divisions, the lower the difference in terms of performance between the thresholds.

It is also interesting that the method that achieves the best recall when the threshold is high is $ALBP_{8,1}$ with a difference of more than 20% with respect to some of the other assessed methods. In summary, for all these reasons, we have fixed the threshold to one.

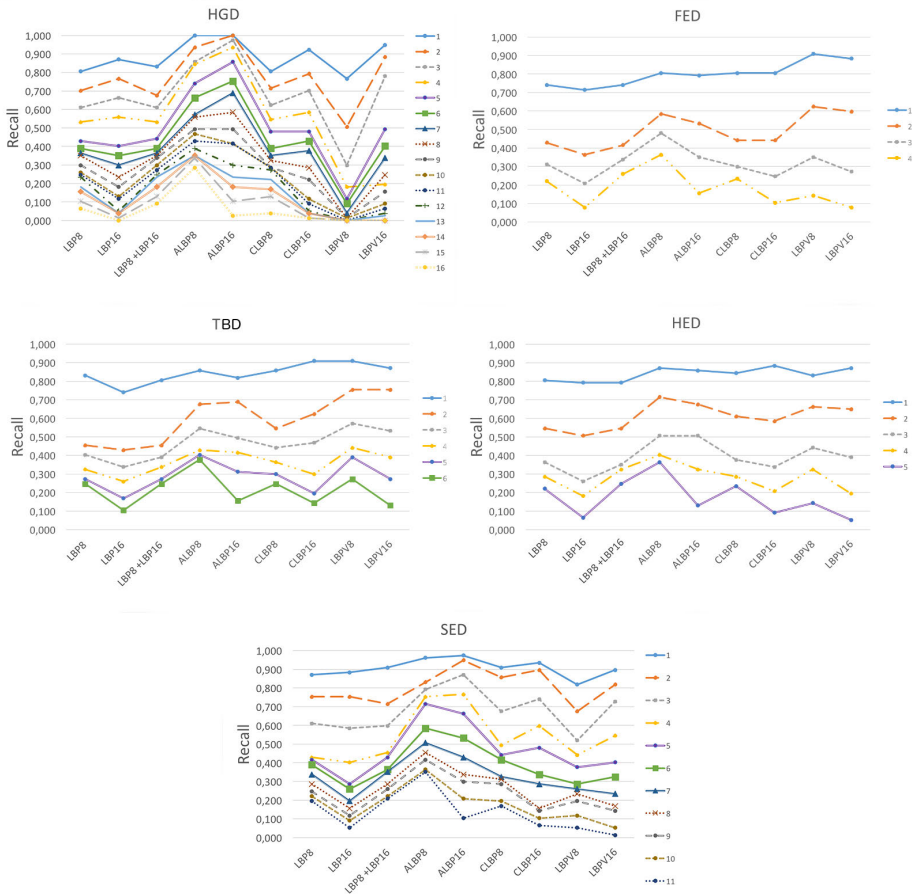


Figure 7.4: Recall values for all wear region configuration methods using different descriptors and threshold values.

7.2.3. Region wear configuration and description evaluation

Figure 7.5 shows the experimental results with threshold value equal to one and all the wear region configuration and descriptor combinations. In this case, four different metrics have been taken into account for studying the classification performance: precision, recall, accuracy and FScore. Usually, FScore is the most representative metric because it takes into account both precision and recall information. As we can observe, the best FScore is achieved using the TBD wear region configuration and CLBP_{16,2} (0.909) and also using the SED configuration and LBP_{8,1}+LBP_{16,2} (0.903), where + represents the concatenation operator.

Another interesting consideration is that although HGD and SED achieves high



Figure 7.5: Results of accuracy, precision, recall and FScore with all the methods and descriptors for the selected threshold value LBP equal to one.

values of recall using ALBP_{8,1} and ALBP_{16,2}, the precision is very low to consider them as a suitable solution.

Taking all the information of Figure 7.4 and Figure 7.5 into account, we can conclude that the best method is the combination composed by SED wear region schema and LBP_{8,1}+LBP_{16,2} due to its high recall (0.909), high FScore (0.903) and high accuracy (90.26%). Moreover, it provides the most stable results in terms of threshold variation which allows the expert more flexibility when varying the

threshold for cases when the precision would be more important than usual. This can be explained because SED method is the one who divides the edge in more number of regions. This allow us to detect more easily small worn patches.

7.3. Conclusions

In conclusion, we have proposed a new method to detect the state of cutting edges (i.e. serviceable or disposable) of inserts in edge profile milling processes. This proposal is an online method based on low cost devices such as the Raspberry Pi and a small monochrome camera (Genie M1280) (See Figure 7.1).

The proposed method divides the cutting edges into several sub-regions (i.e. the wear patches (WP)). Each WP is described using texture descriptors based on Local Binary Pattern (LBP) and classified by means of a Support Vector Machine (SVM) as worn or serviceable. Finally, the decision about whether a cutting edge is serviceable or disposable is based on the number of WP classified as worn.

We have presented several configurations to divide the cutting edge images: Homogeneous grid division (HGD), full edge division (FED), two band division (TBD), half edge division (HED) and small edge division (SED).

We have demonstrated that when at least one of these WP is classified as worn, it is more likely that the insert is disposable.

Using a concatenation of the descriptors $LBP_{8,1}$ and $LBP_{16,2}$ we have achieved an FScore of 0.903 and an accuracy 90.26% dividing the cutting edge following the SED configuration. The main advantage is that our method does not require segmentation of the worn region, which would not be very accurate and subject to segmentation errors. Therefore, the approach presented in this chapter achieves high accuracy and does not require a segmentation stage, what makes feasible to deploy a fully automatic online method to characterise the tool inserts.

8.1. Work summary

One main aim has guided the work presented in this dissertation: Tool wear monitoring using techniques based on computer vision. Real manufacturing enterprises such as TECOI, show high interest in this research line since they require an online and accurate method for determining the status of the wear inserts. In particular, in this Thesis we have focused in description techniques based on shape, contour and texture.

During this dissertation, some new methods have been proposed to deal with this problem: CPDH36, aZIBO, B-ORCHIZ and ShapeFeat. Furthermore, an online and low cost method based on the description of regions of the insert using Local Binary Pattern has been developed. Now, a brief description of these methods is summarized.

CPDH36R is a shape descriptor based on the contour points description. It is an extension of the original CPDH. The authors of CPDH used the EMD classifier in order to make their method invariant to rotation. The main disadvantage of EMD is its high computational cost, thus one of our motivations in developing Contour Points Distribution Histogram Radius using 36 features (CPDH36R) was the necessity of making a new descriptor containing just the most relevant features, and being rotational invariant. CPDH36R also obtains even better hit rates than with the base-line method and it is much more efficient in terms of computational cost.

aZIBO (García-Ordás et al., 2014) and B-ORCHIZ (García-Ordás et al., 2016) are two new proposals based on ZMEG (Zernike Moments Edge Gradient Cooccurrence Matrix) (Anuar et al., 2013), all of them combining global and local descriptors. Experimental results showed that B-ORCHIZ outperformed the others proposals. Unlike ZMEG, we achieved invariance using Zernike moments module up to the tenth order (i.e. 36 values) as global descriptor and we also proposed the IBD (Invariant Boundary Descriptor), composed by IEGCM (Invariant Edge Gradient Co-occurrence Matrix) and BOC (Boundary Orientations Chain), as local ones. We applied our proposal using the images resized to 128×128 pixels and interpolating the binary images between 0 and 255 allowing Zernike moments to be applied on

a wider range of integer values. Moreover, a clustering analysis and a shape study depending on the wear was carried out helping the experts to recognize the different levels of wear thanks to the representative shapes created for each one.

The ShapeFeat descriptor takes into account ten different features extracted of the binary region of each image. This descriptor provides information about the image which is not usually obtained with moment descriptors such as Hu, Flusser, Zernike, etc. that are focused in the boundary information, avoiding interesting information such as the area of the region or the homogeneity of it. Ten features of the binary region were considered: ConvexArea, Eccentricity, Perimeter, EquivDiameter, Extent, FilledArea, FilledImage, MinorAxisLength, MajorAxisLength and Solidity.

An evaluation of B-ORCHIZ and ShapeFeat fusion methods have been carried out, using early fusion, late fusion and intermediate fusion obtaining better results than using each method isolatedly. This combination represents a breakthrough in the tool wear monitoring field.

Finally, we proposed a new online, low cost and fast approach to determine the state and categorize the wear of cutting tools used in edge profile milling processes. It uses devices such as the Raspberry pi or a small monochrome camera (Genie M1280). The method proposed to described the cutting edge region is based on dividing the region in patches avoiding the segmentation stage. We evaluated five alternatives and then, we describe each patch with different texture methods.

In the rest of the chapter, the main conclusions of this Thesis and future work lines are presented.

8.2. General contributions

This dissertation has provided solutions to tool wear monitoring using computer vision algorithms. All methods have been assessed with images obtained directly from a cutting head located in our research lab. The contributions help manufacturing enterprises to categorize the tool wear, saving costs and time. The main contributions derived from this dissertation are:

1. Three new datasets for tool wear monitoring have been created. To do this, a real industry environment for the image acquisition has been recreated contributing to the scientific community making these datasets available in the website of our research group ¹. All datasets have been created taken images of the inserts located in real cutting tools. This results in an easy deployment from laboratory experiments to manufacturing ones.

¹<http://pitia.unileon.es/VARP/galleries>

2. A new method based on CPDH called CPDH36R has demonstrated the high impact of rotational invariance in shape description. Previously, another method called RCPDH was developed trying to get this invariance to rotation. Results demonstrate the good performance of RCPDH using shape retrieval datasets like Kimia, MPEG7 and a subset of it called MPEG2 using the k Nearest Neighbors method. RCPDH achieves the same results than the original CPDH but reducing drastically the computational cost. On the other hand, CPDH36R not only decreases the computational cost but also increases the performance in the classification of the most complex datasets: MPEG7 and MPEG2. In the first case, the improvement of CPDH36R compared with the original method was of a 26.66% in the hit rate metric. Using the subset MPEG2, our proposal also achieved a hit rate that is 7.35% higher than the traditional CPDH.
3. A new method called aZIBO has been proposed to shape retrieval based on moments and rotational invariant features. This method based on ZMEG descriptor has demonstrated a better performance than the original one for shape retrieval datasets. In the case of Kimia99, the improvement is low compared to the high hit rate achieved by ZMEG, an 87.88%. For MPEG2 and MPEG7 datasets, the experiments demonstrated clearly the good performance of our proposal with a 4.85% of improvement with respect to ZMEG in the first dataset and a 3.2% in the second ones. Furthermore, we have evaluated aZIBO for tool wear monitoring comparing its performance against some classic shape descriptor methods like Hu or Flusser. Experiments showed that using the insert edge dataset, aZIBO outperforms all the other methods with a 81.13% of success rate in the two class scenario and 60.37% of success rate in the three class experiment. In the case of the region based dataset the performance achieved is even higher, with a 91.33% of hit rate for two classes (Low-High) and a 90.12% for three (Low-Medium-High) classification, increasing the hit rate in more than a 6.3% compared with the classical methods.
4. B-ORCHIZ is a contour description method which outperforms significantly the results obtained by other contour methods like aZIBO or ZMEG. B-ORCHIZ was evaluated for insert datasets outperforming aZIBO in more than a 3% using a classification process based on a trained Support vector machine model. Additionally, we have performed an automatic clustering analysis based on dendrograms to find the best way to divide the inserts by their wear level. B-ORCHIZ demonstrated to be the best description method evaluated for insert wear classification. Furthermore, a study of the representative shape of wear classes in order to help the experts to easily identify the different levels comparing the new inserts with the

representative shapes is quite innovative. Our experiment showed that B-ORCHIZ descriptor provided the best description for the wear, showing in the different representative shapes more information than the other methods including the manual one. This can be used not only in the study of the wear process but also to train new personal, generate automatic templates for visual assessment, etc.

5. The proposed ShapeFeat is quite innovative in terms of features extracted and performance. This method achieves great results in the insert wear classification using the Insert region dataset with an 88.70% of hit rate for two classes and 80.67% for three classes labelling. Furthermore, the method is formed by just 10 features which saves lot of memory space in the descriptor databases.
6. Fusion of contour and shape features have never been explored for tool wear monitoring. Three different ways of merging were evaluated including early fusion, intermediate fusion and late fusion. Results demonstrate the advantages of combining different kind of features for classification obtaining a 91.44% of hit rate for the binary classification and a 82.90% for the ternary one using late fusion in both cases.
7. A new online method for automatic evaluation system of wear in milling processes has been proposed. It is a low cost system, full automatic and easy to attach to all kind of cutting machines. This system proposed the division of insert images into small patches isolatedly described using texture descriptors. Classification results showed a hit rate of 90.26% with a binary classification. This method has no need of a segmentation step and has an easy implementation on low cost machines.

8.3. Outlook

In this section, we summarise the main research lines that remain open in this application field.

First of all, despite the large amount of shape description methods, there is always a way to improve the existing ones. The final aim must be to implement a method invariant to the maximum number of transformations possible (rotation, illumination, scale,...), that provides representative wear information, good performance as well as efficiency in terms of computational cost.

In the tool wear monitoring field, the last work in the topic up to our knowledge, is our proposal based on dividing the insert images into small patches and describe them using texture descriptors. It is a very promising approach because the system

is a low cost method, full automatic and easy to attach to all kind of cutting machines.

Other way to improve this work is the study of a better illumination system. Our method relies on edge detection and therefore achieving a good contrast is highly important. One possible solution would be to capture the same insert under several lighting positions to combine contour information of all images.

The texture descriptors employed to describe the cutting edge patches, are based on the existing LBP (Local Binary Pattern). As our method depends largely on the texture descriptor used, it would be useful to study and also propose new texture descriptors.

Furthermore, it would be interesting to find the way of fusing local texture description of the patches with the full contour or shape description methods.

Capítulo 9

Conclusiones y perspectiva

9.1. Resumen del trabajo

Un objetivo único ha guiado el trabajo que se presenta en esta Tesis: la monitorización del desgaste de la herramienta de corte utilizando técnicas basadas en visión artificial. Empresas como TECOI tienen un gran interés en esta línea de investigación porque necesitan un método online y preciso para determinar el estado de desgaste de las plaquitas. En particular, en esta Tesis nos centramos en las técnicas de descripción de forma, contorno y textura.

En esta tesis se han propuesto cinco métodos de descripción para abordar este problema. Nuestra primera propuesta, Histograma de Distribución de los Puntos de Contorno, está compuesta por 36 características teniendo en cuenta el radio (CPDH36R). Posteriormente se desarrollaron aZIBO y B-ORCHIZ basados en ZMEG, que combinan descriptores globales y locales, ShapeFeat que se basa en la información de la región de desgaste y finalmente, se propuso un método online, rápido y de bajo coste para la monitorización del desgaste de la herramienta de corte basado en la descripción mediante texturas de parches de la región de desgaste. A continuación resumimos brevemente estos métodos:

CPDH36R es un descriptor de forma que tiene en cuenta la información de los puntos de contorno. Se basa en el método original CPDH. Los autores de CPDH utilizan el clasificador EMD con el fin de hacer su método invariante a la rotación. La principal desventaja de EMD es su lentitud, por tanto, una de nuestras motivaciones para desarrollar CPDH36R, fue la necesidad de hacer un nuevo descriptor que contuviera solo las características más relevantes, y por esta razón obtiene aún mejores tasas de acierto que el método base, y al mismo tiempo es mucho más eficiente en términos de coste computacional.

aZIBO (García-Ordás et al., 2014) y B-ORCHIZ (García-Ordás et al., 2016) son dos nuevas propuestas basadas en ZMEG (Anuar et al., 2013), y los tres combinan descriptores globales y locales. B-ORCHIZ resultó ser la mejor de las dos propuestas. A diferencia de ZMEG, hemos logrado invariancia utilizando el módulo de los momentos de Zernike hasta el décimo orden (es decir, 36 valores) como descriptor global y también propusimos IBD (Invariant Boundary Descriptor) como descriptor local. IBD está compuesto por IEGCM (Invariant Edge Gradient Cooccurrence

Matrix) y BOC (Boundary Orientation Chain). Aplicamos la propuesta utilizando las imágenes redimensionadas a 128×128 píxeles y la interpolación de las imágenes binarias entre 0 y 255 que permite aplicar los momentos de Zernike sobre una gama más amplia de valores enteros. Además, se ha llevado a cabo un análisis de agrupaciones y un estudio de la forma en función del desgaste para ayudar a los expertos a reconocer los diferentes niveles de desgaste gracias a las formas representativas obtenidas para cada una de las clases.

El descriptor ShapeFeat tiene en cuenta diez características diferentes extraídas de la región binaria de cada imagen. Este descriptor proporciona información muy útil acerca de la forma que no se obtiene normalmente con descriptores basados en momentos, como Hu, Flusser, Zernike, etc., ya que no tienen en cuenta datos como el área de la región o la homogeneidad de la misma. En nuestro caso hemos seleccionado diez propiedades. Estas 10 características de la región binaria son: ConvexArea, la excentricidad, el perímetro, EquivDiameter, extensión, FilledArea, FilledImage, MinorAxisLength, MajorAxisLength y solidez.

B-ORCHIZ y ShapeFeat son los mejores descriptores de forma por el momento, así que decidimos combinar ambos con diferentes técnicas de fusión y obtuvimos mejores resultados que usando cada uno de los métodos por separado. Esta combinación resulta ser un avance en el campo de la monitorización de desgaste de herramienta de corte.

Por último, hemos propuesto un nuevo método online, de bajo coste y rápido para determinar el estado y categorizar el desgaste de las herramientas utilizadas en los procesos de fresado. Utiliza dispositivos muy baratos, tales como una Raspberry pi o una pequeña cámara monocromática (Genie M1280). El método propuesto para describir el borde de la región de corte se basa en dividir dicha región en parches o pequeñas regiones evitando el proceso de segmentación. Hemos propuesto cinco alternativas y después de evaluarlas todas y de seleccionar una de ellas, describimos cada parche con diferentes métodos de textura. En el resto de este capítulo se presentan las principales contribuciones de esta Tesis y el trabajo futuro.

9.2. Contribuciones generales

Esta tesis proporciona soluciones para monitorizar el desgaste de las herramientas de corte utilizando diferentes técnicas de descripción de la forma de la región de desgaste. Todos los métodos propuestos fueron evaluados con imágenes obtenidas directamente de un cabezal de corte. Las contribuciones que se derivan de este trabajo son:

1. Hemos creado tres nuevos conjuntos de imágenes para poder evaluar nuestras propuestas de monitorización de desgaste de las herramientas. Para la adquisición de las imágenes se ha recreado un entorno industrial real,

aportando a la comunidad científica tres diferentes conjuntos de imágenes disponibles en la página web del grupo de investigación ¹. Estos conjuntos de datos son muy útiles porque las plaquitas de corte son fotografiadas directamente cuando están colocadas en el cabezal de corte de la herramienta. Eso hace muy fácil el despliegue de los experimentos de laboratorio en los entornos de fabricación.

2. Hemos propuesto un nuevo método basado en CPDH, llamado CPDH36R. Antes de eso, también se desarrolló otro método llamado RCPDH que trata de añadir invariancia a rotación. Los resultados demuestran el buen rendimiento del método propuesto utilizando conjuntos de datos de recuperación de forma como Kimia, MPEG7 y un subconjunto de este último denominado MPEG2 utilizando el método de los k vecinos más cercanos. RCPDH obtiene resultados bastante similares al CPDH original, pero reduce drásticamente el coste computacional. Por otro lado, CPDH36R aumenta el rendimiento en términos de coste computacional y mejora los resultados de la clasificación de los conjuntos de datos más complejos: MPEG7 y MPEG2. En el primer caso, la mejora de CPDH36R en comparación con el método original era de un 26,66 % de tasa de acierto. Usando el subconjunto MPEG2, nuestra propuesta también logró un 7,35 % más de tasa de éxito que con CPDH tradicional.
3. Hemos propuesto un nuevo método de descripción de forma llamado aZIBO basado en momentos y en características invariantes a la rotación. Este método inspirado en el descriptor ZMEG, ha demostrado un mejor rendimiento que el original para conjuntos de imágenes de recuperación de forma. En el caso de Kimia99, la mejora es baja en comparación con la tasa de éxito alcanzado por ZMEG, un 87,88 %. En los conjuntos MPEG2 y MPEG7, los experimentos demostraron claramente el buen rendimiento de nuestra propuesta con un 4,85 % de mejoría con respecto al método ZMEG en los experimentos con los primeros conjuntos de imágenes y un 3,2 % en los segundos. Por otra parte, hemos evaluado aZIBO para la monitorización de desgaste de herramientas comparando su rendimiento frente a algunos de los métodos clásicos como Hu o Flusser. Los experimentos mostraron que utilizando el conjunto de imágenes de borde de las plaquitas, aZIBO supera a todos los otros métodos con un 81,13 % de tasa de éxito en la clasificación de dos clases (bajo-alto desgaste) y 60,37 % de tasa de éxito en el experimento de tres clases (bajo-medio-alto desgaste). En el caso del conjunto de imágenes de regiones de plaquitas, el rendimiento alcanzado es todavía mayor, con un 91,33 % de tasa de acierto en la clasificación binaria (bajo-alto desgaste) y un 90,12 % en el caso de baja-media-alta clasificación, aumentando la tasa de éxito en más de un 6,3 %

¹<http://pitia.unileon.es/VARP/galleries>

en comparación con los métodos clásicos.

4. También hemos presentado una evolución más compleja de aZIBO (García-Ordás et al., 2014) llamada B-ORCHIZ (García-Ordás et al., 2016) para la evaluación del desgaste de las plaquitas. El método, que mejora los resultados para los conjuntos de datos clásicos de recuperación de forma, se evaluó también en los conjuntos de imágenes de plaquitas, superando a los resultados obtenidos con aZIBO en más de un 3% utilizando en este caso un proceso de clasificación basado en un modelo de máquina de vectores de soporte entrenado (SVM). Hemos realizado un análisis de agrupamiento automático basado en dendrogramas para encontrar la mejor manera de dividir las plaquitas de corte teniendo en cuenta su nivel de desgaste. A partir de la información proporcionada por los dendrogramas, B-ORCHIZ demostró ser el mejor método de descripción mostrando la mayor capacidad de discriminación entre clases y cómo de compactas son cada una. Además, hemos llevado a cabo un estudio de la forma representativa de las clases de desgaste con el fin de ayudar a los expertos a identificar fácilmente los diferentes niveles de desgaste comparando las nuevas plaquitas con las formas representativas de cada nivel. Nuestro experimento mostró que el descriptor B-ORCHIZ proporciona la mejor descripción para el desgaste, mostrando en las diferentes formas representativas más información que los demás métodos. Esto es muy interesante, no solo en el estudio del proceso de desgaste, sino también para capacitar a personal nuevo, generar plantillas automáticas, etc.
5. La propuesta ShapeFeat es una innovación en términos de extracción de características y rendimiento. Este método alcanza buenos resultados en la clasificación del desgaste de la plaquita utilizando el conjunto de imágenes de regiones de plaquitas, con el que se obtiene un 88,70% de tasa de acierto en la clasificación binaria (bajo-alto desgaste) y 80,67% para el etiquetado de tres clases (bajo-medio-alto desgaste). Además, el método está formado por solo 10 características, lo cual ahorra mucho espacio de memoria en las bases de datos de descriptores.
6. Nunca antes se había explorado la fusión de descriptores de forma con descriptores de contorno para monitorizar el desgaste de herramientas de corte. Nosotros hemos evaluado tres métodos de fusión de descriptores: fusión temprana, intermedia y tardía. Los resultados muestran las ventajas de combinar ShapeFeat y BORCHIZ, obteniendo un 91,44% de tasa de acierto para la clasificación binaria y un 82,90% para la ternaria, utilizando fusión tardía en ambos casos.
7. Hemos presentado un nuevo método online para la evaluación automática de desgaste en los procesos de fresado. El sistema es de coste muy reducido,

totalmente automático y fácil de conectar a todo tipo de máquinas de corte. Este sistema propone una división de la región de desgaste de la plaquita de corte en pequeños parches y la descripción de cada uno de los parches aisladamente. En la etapa de clasificación mostró una tasa de acierto del 90,26 % para la clasificación binaria, que es similar al mejor resultado obtenido por el método de cotransducción, pero en cuyo caso, el proceso no es completamente automático como lo es en esta propuesta.

9.3. **Perspectiva**

En esta sección, se resumen las principales líneas de investigación que permanecen abiertas con relación al trabajo desarrollado durante esta tesis.

En primer lugar, a pesar de la gran cantidad de métodos de descripción de forma, siempre hay una manera de mejorar los existentes. Deberíamos poder contar con un método de descripción invariante al número máximo de transformaciones posibles (rotación, iluminación, escala, ...), que obtenga buenos resultados y que sea también eficiente y con un bajo coste computacional.

En el campo de la monitorización de desgaste de herramientas, el último trabajo en el tema hasta donde sabemos, es nuestra propuesta basada en la división de las imágenes de plaquitas en pequeños parches aislados y su posterior descripción usando métodos de textura. Es un enfoque muy prometedor debido a que el sistema es de bajo coste, totalmente automático y fácil de conectar a todo tipo de máquinas de corte.

Las formas de mejorar este trabajo pueden ser, por un lado, estudiar un mejor sistema de iluminación. Nuestro método se basa en la detección de bordes y por lo tanto, el hecho de lograr un buen contraste es muy importante. Una posible solución sería capturar la misma plaquita bajo varias posiciones de iluminación para combinar información de contorno de todas las imágenes.

Por otra parte, los descriptores de textura empleados para describir los parches del filo de corte, se basan en el LBP existente (Local Binary Pattern). Ya que nuestro método depende en gran medida del descriptor de textura usado, sería muy interesante estudiar y proponer nuevos descriptores de textura.

Bibliography

- Ahonen, T., Hadid, A. and Pietikainen, M.: 2006, Face description with local binary patterns: Application to face recognition, *IEEE transactions on pattern analysis and machine intelligence* .
- Ai, T., Cheng, X., Liu, P. and Yang, M.: 2013, A shape analysis and template matching of building features by the fourier transform method, *Computers, Environment and Urban Systems* **41**, 219 – 233.
- Andaló, F., Miranda, P., da S. Torres, R. and Falcao, A.: 2010, Shape feature extraction and description based on tensor scale, *Pattern Recognition* **43**, 26–36.
- Antić, A., Popović, B., Krstanović, L., Obradović, R. and Milosević, M.: 2018, Novel texture-based descriptors for tool wear condition monitoring, *Mechanical Systems and Signal Processing* **98**, 1 – 15.
- Anuar, F. M., Setchi, R. and kun Lai, Y.: 2013, Trademark image retrieval using an integrated shape descriptor, *Expert Systems with Applications* **40**(1), 105 – 121.
- Azmi, A.: 2015, Monitoring of tool wear using measured machining forces and neuro-fuzzy modelling approaches during machining of GFRP composites, *Advances in Engineering Software* **82**, 53 – 64.
- Bai, X., Wang, B., Yao, C., Liu, W. and Tu, Z.: 2012, Co-transduction for shape retrieval, *IEEE Transactions on Image Processing* .
- Barreiro, J., Castejón, M., Alegre, E. and Hernández, L.: 2008, Use of descriptors based on moments from digital images for tool wear monitoring, *International Journal of Machine Tools and Manufacture* **48**(9), 1005 – 1013.
- Belongie, S., J. Malik and Puzicha, J.: 2002, Shape matching and object recognition using shape contexts, *IEEE Transactions on Pattern Analysis and Machine Intelligence* .
- Bhuiyan, M., Choudhury, I., Dahari, M., Nukman, Y. and Dawal, S.: 2016, Application of acoustic emission sensor to investigate the frequency of tool wear and plastic deformation in tool condition monitoring, *Measurement* **92**, 208 – 217.

- Biswas, S., Aggarwal, G. and Chellappa, R.: 2010, An efficient and robust algorithm for shape indexing and retrieval, *Multimedia, IEEE Transactions on* **12**(5), 372–385.
- Blenkinsopp, R., Harland, A., Price, D., Lucas, T. and Roberts, J.: 2012, A method to measure dynamic dorsal foot surface shape and deformation during linear running using digital image correlation, *Procedia Engineering* **34**(0), 266 – 271. Engineering of sport conference 2012.
- Blum, A. and Mitchell, T.: 1998, Combining labeled and unlabeled data with co-training, *Proceedings of the Eleventh Annual Conference on Computational Learning Theory, COLT' 98*, ACM, New York, NY, USA, pp. 92–100.
- Bolourchi, P., Demirel, H. and Uysal, S.: 2017, Target recognition in sar images using radial chebyshev moments, *Signal, Image and Video Processing* pp. 1–8.
- Bostrom, H.: 2007, Feature vs. classifier fusion for predictive data mining a case study in pesticide classification.
- Bovic Kilundu, Pierre Dehombreux, X. C.: 2011, Tool wear monitoring by machine learning techniques and singular spectrum analysis, *Mechanical Systems and Signal Processing* **25**, 400–415.
- Cambuim, L. F., Macieira, R. M., Neto, F. M., Barros, E., Ludermir, T. B. and Zanchettin, C.: 2016, An efficient static gesture recognizer embedded system based on elm pattern recognition algorithm, *Journal of Systems Architecture* **68**, 1 – 16.
- Ceau, G., Popovici, V. and Croitoru, S.: 2010, Researches about the temperature of the cutting edge in turning of unalloyed steel.
- Choi, D., Kwon, T. W. and Chu, N. C.: 1999, Real-time monitoring of tool fracture in turning using sensor fusion, *The International Journal of Advanced Manufacturing Technology* **15**(5), 305–310.
- Chryssolouris, G., Mourtzis, D., Beno, T. and Hulling, U.: 2012, 45th cirp conference on manufacturing systems 2012 measurement of cutting edge temperature in drilling, *Procedia CIRP* **3**, 531 – 536.
- Condorovici, R. G., Florea, C. and Vertan, C.: 2015, Automatically classifying paintings with perceptual inspired descriptors, *Journal of Visual Communication and Image Representation* **26**, 222 – 230.
- Connors, R. W. and Harlow, C. A.: 1980, A theoretical comparison of texture algorithms, *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI* .
- Curic, V., Lindblad, J., Sladoje, N., Sarve, H. and Borgefors, G.: 2014, A new set distance and its application to shape registration, *Pattern Analysis and Applications* **17**(1), 141–152.
- Czerniawski, T., Nahangi, M., Haas, C. and Walbridge, S.: 2016, Pipe spool recognition in cluttered point clouds using a curvature-based shape descriptor, *Automation in Construction* **71, Part 2**, 346 – 358.

- da Silva, R. H. L., da Silva, M. B. and Hassui, A.: 2016, A probabilistic neural network applied in monitoring tool wear in the end milling operation via acoustic emission and cutting power signals, *Machining Science and Technology* .
- Dalal, N. and Triggs, B.: 2005, Histograms of oriented gradients for human detection, 2005 *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, Vol. 1, pp. 886–893 vol. 1.
- Datta, A., Dutta, S., Pal, S. and Sen, R.: 2013, Progressive cutting tool wear detection from machined surface images using voronoi tessellation method, *Journal of Materials Processing Technology* **213**(12), 2339 – 2349.
- Dimla, D.: 1999, Tool wear monitoring using cutting force measurements, *Advances in Manufacturing Technology* .
- Dimla, E. and Dimla, S.: 2000, Sensor signals for tool-wear monitoring in metal cutting operations—a review of methods, *International Journal of Machine Tools and Manufacture* **40**(8), 1073 – 1098.
- Dominguez-Caballero, J. A., Manson, G. A. and Marshall, M. B.: 2016, Tool condition monitoring of ceramic inserted tools in high speed machining through image processing, *International Journal of Mechanical, Aerospace, Industrial, Mechatronic and Manufacturing Engineering* **10**(8), 1413 – 1420.
- Doost, H. R. E. and Amirani, M. C.: 2013, Texture classification with local binary pattern based on continues wavelet transformation, *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering* .
- Dutta, S., Datta, A., Chakladar, N. D., Pal, S., Mukhopadhyay, S. and Sen, R.: 2012, Detection of tool condition from the turned surface images using an accurate grey level co-occurrence technique, *Precision Engineering* **36**(3), 458 – 466.
- Dutta, S., Pal, S. K. and Sen, R.: 2016a, On-machine tool prediction of flank wear from machined surface images using texture analyses and support vector regression, *Precision Engineering* **43**, 34 – 42.
- Dutta, S., Pal, S. K. and Sen, R.: 2016b, Progressive tool flank wear monitoring by applying discrete wavelet transform on turned surface images, *Measurement* **77**, 388 – 401.
- Dutta, S., Pal, S., Mukhopadhyay, S. and Sen, R.: 2013, Application of digital image processing in tool condition monitoring: A review, *CIRP Journal of Manufacturing Science and Technology* **6**(3), 212 – 232.
- El-ghazal, A., Basir, O. and Belkasim, S.: 2012, Invariant curvature-based fourier shape descriptors, *Journal of Visual Communication and Image Representation* **23**(4), 622 – 633.
- Fernández-Robles, L., Azzopardi, G., Alegre, E. and Petkov, N.: 2015a, *Cutting Edge Localisation in an Edge Profile Milling Head*, Springer International Publishing, Cham, pp. 336–347.

- Fernández-Robles, L., Azzopardi, G., Alegre, E. and Petkov, N.: 2015b, *Cutting Edge Localisation in an Edge Profile Milling Head*, Springer International Publishing, Cham, chapter Computer Analysis of Images and Patterns: 16th International Conference, CAIP 2015, Valletta, Malta, September 2-4, 2015, Proceedings, Part II, pp. 336–347.
- Fishbaugh, J., Durrleman, S., Prastawa, M. and Gerig, G.: 2017, Geodesic shape regression with multiple geometries and sparse parameters, *Medical Image Analysis* **39**, 1 – 17.
- García-Olalla, O., Alegre, E., Fernández-Robles, L., Malm, P. and Bengtsson, E.: 2015, Acrosome integrity assessment of boar spermatozoa images using an early fusion of texture and contour descriptors, *Computer Methods and Programs in Biomedicine* **120**(1), 49–64.
- García-Olalla, O., Fernández-Robles, L., Fidalgo, E., González-Castro, V. and Alegre, E.: 2017, *Evaluation of the State of Cutting Tools According to Its Texture Using LOSIB and LBP Variants*, Springer International Publishing, Cham, pp. 217–228.
- García-Ordás, M., Alegre, E., González-Castro, V. and García-Ordás, D.: 2014, azibo: A new descriptor based in shape moments and rotational invariant features, *Pattern Recognition (ICPR), 2014 22nd International Conference on*, pp. 2395–2400.
- García-Ordás, M. T., Alegre, E., González-Castro, V. and Alaiz-Rodríguez, R.: 2016, A computer vision approach to analyse and classify tool wear level in milling processes using shape descriptors and machine learning techniques, *International Journal of Advanced Manufacturing Technology* .
- Guo, Z., Zhang, L. and Zhang, D.: 2010a, A completed modeling of local binary pattern operator for texture classification, *Image Processing, IEEE Transactions on* .
- Guo, Z., Zhang, L. and Zhang, D.: 2010b, Rotation invariant texture classification using lbp variance (lbpv) with global matching, *Pattern Recognition* .
- Guo, Z., Zhang, L., Zhang, D. and Zhang, S.: 2010, Rotation invariant texture classification using adaptive lbp with directional statistical features, *Image Processing (ICIP), 2010 17th IEEE International Conference on* .
- Hadizadeh, H.: 2015, Multi-resolution local gabor wavelets binary patterns for gray-scale texture description, *Pattern Recognition Letters* **65**, 163 – 169.
- Halidou, A., You, X., Hamidine, M., Etoundi, R. A., Diakite, L. H. and Souleimanou: 2014, Fast pedestrian detection based on region of interest and multi-block local binary pattern descriptors, *Computers & Electrical Engineering* **40**(8), 375 – 389.
- Hao, S., Jiang, J., Guo, Y. and Li, H.: 2013, Active learning based intervertebral disk classification combining shape and texture similarities, *Neurocomputing* **101**(0), 252 – 257.
- Haralick, R. M., Shanmugam, K. and Dinstein, I.: 1973, Textural features for image classification, *IEEE Trans. on Systems, Man and Cybernetics* .
- Hoang, V.-D., Le, M.-H. and Jo, K.-H.: 2014, Hybrid cascade boosting machine using variant scale blocks based {HOG} features for pedestrian detection, *Neurocomputing* **135**, 357 – 366.

- Hu, M.-K.: 1962, Visual pattern recognition by moment invariants, *IRE Transactions on Information Theory* **8**(2), 179–187.
- Hu, R.-X., Jia, W., Ling, H., Zhao, Y. and Gui, J.: 2014, Angular pattern and binary angular pattern for shape retrieval, *Image Processing, IEEE Transactions on* **23**(3), 1118–1127.
- Hu, R.-X., Jia, W., Zhao, Y. and Gui, J.: 2012, Perceptually motivated morphological strategies for shape retrieval, *Pattern Recognition* **45**(9), 3222 – 3230. Best Papers of Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA'2011).
- Kassim, A., Mannan, M. and Zhu, M.: 2007, Texture analysis methods for tool condition monitoring, *Image and Vision Computing* **25**, 1080–1090.
- Kaya, B., Oysu, C. and Ertunc, H. M.: 2011, Force-torque based on-line tool wear estimation system for CNC milling of inconel 718 using neural networks, *Advances in Engineering Software* **42**(3), 76 – 84.
- Khotanzad, A. and Hong, Y. H.: 1990, Invariant image recognition by zernike moments, *IEEE Trans. Pattern Anal. Mach. Intell.* **12**(5), 489–497.
- Kim, D. H. and Park, S.-K.: 2014, Texture-less object recognition using contour fragment-based features with bisected local regions, *Consumer Electronics (ICCE), 2014 IEEE International Conference on*, pp. 388–389.
- Kurada, S. and Bradley, C.: 1997, A review of machine vision sensors for tool condition monitoring, *Computers in Industry* **34**(1), 55 – 72.
- Laiche, N., Larabi, S., Ladraa, F. and Khadraoui, A.: 2014, Curve normalization for shape retrieval, *Signal Processing: Image Communication* **29**(4), 556 – 571.
- Li, L. and An, Q.: 2016, An in-depth study of tool wear monitoring technique based on image segmentation and texture analysis, *Measurement* **79**, 44 – 52.
- Li, X.: 2002, A brief review: acoustic emission method for tool wear monitoring during turning, *International Journal of Machine Tools and Manufacture* **42**(2), 157 – 165.
- Li, X., Ruan, Q., Jin, Y., An, G. and Zhao, R.: 2015, Fully automatic 3d facial expression recognition using polytypic multi-block local binary patterns, *Signal Processing* **108**, 297 – 308.
- Loizou, J., Tian, W., Robertson, J. and Camelio, J.: 2015, Automated wear characterization for broaching tools based on machine vision systems, *Journal of Manufacturing Systems* **37**, Part 2, 558 – 563.
- Madrid-Cuevas, F., Aguilera-Aguilera, E., Carmona-Poyato, A., Muñoz-Salinas, R., Medina-Carnicer, R. and Fernández-García, N.: 2016, An efficient unsupervised method for obtaining polygonal approximations of closed digital planar curves, *Journal of Visual Communication and Image Representation* **39**, 152 – 163.
- Malekian, M., Park, S. S. and Jun, M. B.: 2009, Tool wear monitoring of micro-milling operations, *Journal of Materials Processing Technology* **209**(10), 4903 – 4914.

- Masoumi, M., Li, C. and Hamza, A. B.: 2016, A spectral graph wavelet approach for nonrigid 3d shape retrieval, *Pattern Recognition Letters* **83, Part 3**, 339 – 348. Efficient Shape Representation, Matching, Ranking, and its Applications.
- Mingqiang, Y., Idiyo, K. K. and Joseph, R.: 2008, A Survey of Shape Feature Extraction Techniques, *Pattern Recognition, Peng-Yeng Yin* pp. 43–90.
- Mohanaiah, P., Sathyanarayana, P. and GuruKumar, L.: 2013, Image texture feature extraction using glcm approach, *International Journal of Scientific and Research Publications* .
- Mukanova, A., Gao, Q. and Hu, G.: 2014, N-gram based image representation and classification using perceptual shape features, *Computer and Robot Vision (CRV), 2014 Canadian Conference on*, pp. 349–356.
- Nasreddine, K., Benzinou, A. and Fablet, R.: 2010, Variational shape matching for shape classification and retrieval, *Pattern Recognition Letters* **31(12)**, 1650 – 1657. Pattern Recognition of Non-Speech Audio.
- Nouri, M., K.Fussell, B., L.Ziniti, B. and Linder, E.: 2015, Real-time tool wear monitoring in milling using a cutting condition independent method, *International Journal of Machine Tools & Manufacture* **89**, 1–13.
- Ojala, T. and Pietikainen, M.: 1996, A comparative study of texture measures with classification based on feature distributions, *Pattern Recognition* .
- Ojala, T. and Pietikainen, M.: 2002, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns., *IEEE Transactions on Pattern Analysis and Machine Intelligence* .
- Pedrosa, G. V., Batista, M. A. and Barcelos, C. A.: 2013, Image feature descriptor based on shape salience points, *Neurocomputing* **120**, 156–163.
- Pradeep Kumar, T., Veera Prasad Reddy, M. and Bora, P. K.: 2017, *Leaf Identification Using Shape and Texture Features*, Springer Singapore, Singapore, pp. 531–541.
- Prasad, B. S. and Babu, M. P.: 2017, Correlation between vibration amplitude and tool wear in turning: Numerical and experimental analysis, *Engineering Science and Technology, an International Journal* **20(1)**, 197 – 211.
- Proen, H. and Santos, G.: 2012, Fusing color and shape descriptors in the recognition of degraded iris images acquired at visible wavelengths, *Computer Vision and Image Understanding* **116(2)**, 167 – 178.
- Rao, K. V., Murthy, B. and Rao, N. M.: 2014, Prediction of cutting tool wear, surface roughness and vibration of work piece in boring of AISI 316 steel with artificial neural network, *Measurement* **51(0)**, 63 – 70.
- Ruta, D. and Gabrys, B.: 2000, An overview of classifier fusion methods.
- Scheffer, C. and Heyns, P.: 2001, Wear monitoring in turning operations using vibration and strain measurements, *Mechanical Systems and Signal Processing* **15(6)**, 1185 – 1202.

- Segreto, T., Simeone, A. and Teti, R.: 2013, Multiple sensor monitoring in nickel alloy turning for tool wear assessment via sensor fusion., *8th CIRP Conference on Intelligent Computation in Manufacturing Engineering* pp. 85–90.
- Shahsavari, V., Chouinard, L. and Bastien, J.: 2017, Wavelet-based analysis of mode shapes for statistical detection and localization of damage in beams using likelihood ratio test, *Engineering Structures* **132**, 494 – 507.
- Shakshuki, E., Hammoudi, K., Benhabiles, H., Kasraoui, M., Ajam, N., Dornaika, F., Radhakrishnan, K., Bandi, K., Cai, Q. and Liu, S.: 2015, The 6th international conference on ambient systems, networks and technologies (ant-2015), the 5th international conference on sustainable energy information technology (seit-2015) developing vision-based and cooperative vehicular embedded systems for enhancing road monitoring services, *Procedia Computer Science* **52**, 389 – 395.
- Shu, X., Pan, L. and Wu, X.-J.: 2015, Multi-scale contour flexibility shape signature for fourier descriptor, *Journal of Visual Communication and Image Representation* **26**, 161 – 167.
- Shu, X. and Wu, X.-J.: 2011a, A novel contour descriptor for 2d shape matching and its application to image retrieval, *Image and Vision Computing* **29**(4), 286 – 294.
- Shu, X. and Wu, X. J.: 2011b, A novel contour descriptor for 2d shape matching and its application to image retrieval, *Image and Vision Computing* .
- Singh, C. and Pooja: 2012, An effective image retrieval using the fusion of global and local transforms based features, *Optics & Laser Technology* **44**, 2249–2259.
- Sokic, E. and Konjicija, S.: 2016, Phase preserving fourier descriptor for shape-based image retrieval, *Signal Processing: Image Communication* **40**, 82 – 96.
- Sun, H., Gao, D., Zhao, Z. and Tang, X.: 2017, An approach to in-process surface texture condition monitoring, *Robotics and Computer-Integrated Manufacturing* **48**, 254 – 262.
- Tu, G. J., Hansen, M. K., Kryger, P. and Ahrendt, P.: 2016, Automatic behaviour analysis system for honeybees using computer vision, *Computers and Electronics in Agriculture* **122**, 10 – 18.
- Wang, C., Ming, W. and Chen, M.: 2016, Milling tool's flank wear prediction by temperature dependent wear mechanism determination when machining inconel 182 overlays, *Tribology International* **104**, 140 – 156.
- Wang, G., Yang, Y., Xie, Q. and Zhang, Y.: 2014, Force based tool wear monitoring system for milling process based on relevance vector machine, *Advances in Engineering Software* **71**(1), 46 – 51.
- Wang, J., Xie, J., Zhao, R., Zhang, L. and Duan, L.: 2017, Multisensory fusion based virtual tool wear sensing for ubiquitous manufacturing, *Robotics and Computer-Integrated Manufacturing* **45**, 47 – 58. Special Issue on Ubiquitous Manufacturing (UbiM).
- Wang, Q., Chang, L., Zhou, M., Li, Q., Liu, H. and Guo, F.: 2016, A spectral and morphologic method for white blood cell classification, *Optics & Laser Technology* **84**, 144 – 148.

- Wang, X., Feng, B., Bai, X., Liu, W. and Latecki, L. J.: 2014, Bag of contour fragments for robust shape classification, *Pattern Recognition* **47**(6), 2116 – 2125.
- Ward, J. H.: 1963, Hierarchical grouping to optimize an objective function, *Journal of the American Statistical Association* **58**(301), 236–244.
- Wei, C.-H., Li, Y., Chau, W.-Y. and Li, C.-T.: 2009, Trademark image retrieval using synthetic features for describing global shape and interior structure, *Pattern Recognition* **42**(3), 386 – 394.
- Wu, D., Dang, D. and Wang, J.: 2016, Recognition of students combining features of zernike moment and optical flow, *2016 2nd IEEE International Conference on Computer and Communications (ICCC)*, pp. 676–679.
- Yasseen, Z., Verroust-Blondet, A. and Nasri, A.: 2016, Shape matching by part alignment using extended chordal axis transform, *Pattern Recognition* **57**, 115 – 135.
- Yuan, Z., Li, F., Zhang, P. and Chen, B.: 2014, Description of shape characteristics through fourier and wavelet analysis, *Chinese Journal of Aeronautics* **27**(1), 160 – 168.
- Zagoris, K., Ergina, K. and Papamarkos, N.: 2011, Image retrieval systems based on compact shape descriptor and relevance feedback information, *Journal of Visual Communication and Image Representation* **22**(5), 378 – 390.
- Zayed, N. and Elnemr, H. A.: 2015, Statistical analysis of haralick features to discriminate lung abnormalities, *International Journal of Biomedical Imaging* .
- Zhou, J. H., Pang, C. K., Zhong, Z. W. and Lewis, F. L.: 2011, Tool wear monitoring using acoustic emissions by dominant-feature identification, *IEEE Transactions on Instrumentation and Measurement* **60**(2), 547–559.
- Zhu, C., Bichot, C.-E. and Chen, L.: 2013, Image region description using orthogonal combination of local binary patterns enhanced with color information, *Pattern Recognition* **46**, 1949–1963.

Annex A: Research activities

Publications related with this manuscript

- María Teresa García-Ordás, Enrique Alegre Gutiérrez, Rocío Alaiz-Rodríguez and Víctor González-Castro. "Combining shape and contour features to improve tool wear monitoring in milling processes." *International Journal of Production Research*, On revision. 2017
- María Teresa García-Ordás, Enrique Alegre, Rocío Alaiz-Rodríguez and Víctor González-Castro, "Tool wear monitoring using an online, automatic and low cost system based on local texture" *Mechanical Systems and Signal Processing*, On revision. 2017
- María Teresa García-Ordás, Enrique Alegre, Víctor González-Castro and Rocío Alaiz-Rodríguez, "A computer vision approach to analyse and classify tool wear level in milling processes using shape descriptors and machine learning techniques", *International Journal of Advanced Manufacturing Technology*, 2016. Volume 90 (5), pp. 1947-1961.
- García-Ordás, M.T. and Alegre, E. and González-Castro, V. and García-Ordás, D., "aZIBO: A New Descriptor Based in Shape Moments and Rotational Invariant Features". 22nd International Conference on Pattern Recognition (ICPR), Stockholm, 2014, Proceedings. pp. 2395-2400.
- García-Ordás, M T., Enrique Alegre, O. García-Olalla, and D. García-Ordás, "Evaluation of Different Metrics for Shape Based Image Retrieval Using a New Contour Points Descriptor". *Similarity Search and Applications: 6th International Conference, SISAP 2013, A Coruña, Spain, October 2-4, 2013, Proceedings*. pp 141-150.
- García-Ordás, M T., O. García-Olalla, L. Fernández-Robles, D. García-Ordás, and E. Alegre, "Rotation Invariant Contour Points Descriptor Histogram for shape based image retrieval". *XXXIV Jornadas de Automática, 2013 In proceedings*.

Patents related with this manuscript

- "Sistema y procedimiento para determinar el desgaste en plaquitas de corte utilizadas en operaciones de fresado mediante clasificación de descriptores de forma en imágenes digitales". University of León, Marzo 2017.

- "Sistema y procedimiento para determinar el desgaste en plaquitas de corte utilizadas en operaciones de fresado." University of León, November 2016.

Other publications

- Fidalgo, E., L. Fernández-Robles, M T. García-Ordás, O. García-Olalla, and E. Alegre, "Evaluation of shape and color descriptors by using bag of Words techniques with one vs all classification". XXXIV Jornadas de Automática, Terrassa , 2013.
- García-Ordás, D., Alegre, E., García-Ordás, M T., García-Olalla, O. and Fernández-Robles, L. "Robustness to rotation in perceptual hashing methods via dominant orientation ". XXXIV Jornadas de Automática, Terrassa , 2013.
- García-Olalla, O., E. Alegre, M T. García-Ordás, and L. Fernández-Robles. " Evaluation of LBP Variants Using Several Metrics and kNN Classifiers" International Conference on Similarity Search and Applications (SISAP), 2013.
- Fernández-Robles, L., E. Alegre, M T. García-Ordás, D. García-Ordás, and E. Fidalgo, "Combining SURF with global texture descriptors for classifying board sperm". XXXIV Jornadas de Automática, Terrassa , 2013.
- García-Olalla, O., E. Alegre, L. Fernández-Robles, M T. García-Ordás, and D. García-Ordás, "Adaptive local binary pattern with oriented standard deviation (ALBPS) for texture Classification", EURASIP Journal on Image and Video Processing, 2013.
- García-Ordás, D., L. Fernández-Robles, E. Alegre, M T. García-Ordás, and O. García-Olalla, "Automatic tampering detection in spliced images with different compression levels". IbPRIA 2013 6th Iberian Conference on Pattern Recognition and Image Analysis, 2013.
- García-Olalla, O., M T. García-Ordás, L. Fernández-Robles, D. García-Ordás, and E. Alegre, "Comparison of different Local Binary Pattern variants for material recognition using KTH-TIPS 2a dataset ". XXXIV Jornadas de Automática, Terrassa , 2013.
- García-Ordás, M T., L. Fernández-Robles, O. García-Olalla, D. García-Ordás, and E. Alegre, " Boar spermatozoa classification using local invariant features and bag of words". XXXIII Jornadas de Automática, Vigo, Spain, 2012.
- García-Olalla, O., M T. García-Ordás, D. García-Ordás, L. Fernández-Robles, and E. Alegre, "Vitality assessment of boar sperm using N Concentric Squares resized and Local binary pattern in gray scale images". XXXIII Jornadas de Automática, Vigo, Spain, 2012.
- Fernández-Robles, L., O. García-Olalla, M T. García-Ordás, D. García-Ordás, and E. Alegre, "SVM approach to classify boar acrosome integrity of a multi-features SURF description". XXXIII Jornadas de Automática, Vigo, Spain, 2012.
- Fidalgo, E., J. de Pedro, L. Fernández-Robles, M T. García-Ordás, and E. Alegre, "Evaluation of segmentation methods applied to intact and damaged boar spermatozoon heads". XXXIII Jornadas de Automática, Vigo, Spain, 2012.

-
- V. González-Castro, E. Alegre, Oscar García-Olalla, María Teresa García-Ordás, "Adaptive pattern spectrum image description using Euclidean and Geodesic distance without training for texture classification" .IET Computer Vision, 2012.
 - E. Alegre, Víctor González-Castro, Rocío Aláiz-Rodríguez, María Teresa García-Ordás, "Texture and momento-based classification of the acrosome integrity of boar spermatozoa images". Computer Methods and Programs in Biomedicine, 2012.
 - Morala-Arguello, P., J. Barreiro, E. Alegre, M T. García-Ordás, O. García-Olalla, and D. González-Madruga, "Reliability of Monitoring Signals for Estimation of Surface Roughness in Metallic Turned Parts". Advanced materials research, 2012.
 - Oscar García-Olalla, Enrique Alegre, Laura Fernández-Robles and María Teresa García-Ordás, "Vitality assessment of boar sperm using an adaptive LBP based on oriented deviation". Asian Conference in Computer Vision (ACCV). Local binary pattern Workshop. Daejeon, Korea, 2012.
 - Víctor González-Castro, Enrique Alegre, Oscar García-Olalla, Diego García-Ordás, María Teresa García-Ordás, and Laura Fernández-Robles, " Curvelet-based texture description to classify intact and damaged boar spermatozoa" International Conference on Image Analysis and Recognition ICIAR. Aveiro, Portugal., 2012.
 - M.T García-Ordás, L.Fernández-Robles, O.García-Olalla, Diego García-Ordás, E.Alegre, " Words recognition using methods of word shape coding". XXXII Jornadas de Automática, Sevilla, España, 2011.
 - Enrique Alegre, María Teresa García, Víctor González-Castro, S. Karthikeyan, "Vitality assessment of boar sperm using NCSR texture descriptor in digital images", Lecture Notes in Computer Science. Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA), Las Palmas de Gran Canaria, 2011.
 - L.Fernández-Robles, M.T García-Ordás, Diego García-Ordás, O.García-Olalla, E.Alegre "Acrosome evaluation of spermatozoa cells using SIFT and classical texture Descriptors". XXXII Jornadas de Automática, Sevilla, España, 2011.
 - Enrique Alegre Gutiérrez, María Teresa García-Ordás, Oscar García-Olalla, Víctor González-Castro, "Cómo escribir documentos científicos. Elaboración de tesis doctorales y artículos de ingeniería". Libro. Editorial: CERSA. Lugar de publicación: Madrid, 2011.
 - O.García-Olalla, Diego García-Ordás, M.T García-Ordás, L. Fernández-Robles, E.Alegre, "Adaptative filters evaluation for sharpness enhancement and noise removal". XXXII Jornadas de Automática, Sevilla, España, 2011.
 - Suárez, S A., J. Barreiro, E. Alegre, M T. García-Ordás, and O. García-Olalla, "Material surface characterization using laws descriptors" . The 22st DAAAM International Symposium, Viena, 2011.
 - Diego García-Ordás, O.García-Ordás, L. Fernández-Robles, M.T García-Ordás, E.Alegre, " Video segmentation combining depth maps and intensity images". XXXII Jornadas de Automática, Sevilla, España, 2011.

- P.Morala-Arguello, J. Barreiro, E. Alegre, M. García-Ordás, Oscar García-Olalla, “ Surface roughness prediction from combination of cutting forces, turning vibrations and machining conditions using artificial neural networks”. 4th Manufacturing Engineering Society International Conference (MESIC). Cádiz, España, 2011.
- L. Fernández-Robles, V. González-Castro, O. García-Olalla, M.T. García-Ordás, E. Alegre, “ A local invariant features approach for classifying acrosome integrity in boar spermatozoa”. III Ecomas thematic conference on computational vision and medical image processing. Algarve, Portugal, 2011.
- Víctor González-Castro, Enrique Alegre, Sir Alexci Suárez, Oscar García-Olalla, María Teresa García, “Adaptative texture description for semen vitality assessment”. Seminario Internacional de Procesamiento y Análisis de Imágenes Médicas (International seminar of medical images processing and analysis) (SIPAIM). Bogotá, Colombia, 2010.
- Enrique Alegre, Maribel González, Víctor González-Castro, Tomás Alonso, María Teresa García, Oscar García-Olalla, “Evaluation of mother Wavelet functions with statistical texture descriptors to classify boar sperm acrosome”. Seminario Internacional de Procesamiento y Análisis de Imágenes Médicas (International seminar of medical images processing and analysis) (SIPAIM). Bogotá, Colombia, 2010.

Co-author in the following inventions:

- Patent: Artificial vision procedure for the detection of proximal cytoplasmic droplets in spermatozoa, 2014.
- Patent: Artificial vision procedure for the detection of spermatozoa with curled tails, 2013.
- Intellectual property: Detection of proximal droplets in tails of spermatozoa using artificial vision techniques, 2011.
- Intellectual property: Detection of distal droplets in tails of spermatozoa using artificial vision techniques, 2011.
- Intellectual property: Classification of spermatozoa according to the state of their acrosomes using artificial vision techniques, 2011.
- Intellectual property: Detection of heads of spermatozoa with whip tails using artificial vision techniques, 2011.

Research projects

- Computer vision systems for life prediction of cutters for machining in severe conditions using fusion signals based classification. Spanish Ministry of Science and Innovation.

Financing entity: MICINN

Duration from: 01/01/2013 to: 31/12/2015

Main researcher: Enrique Alegre Gutiérrez

Quantity: 84700 €

-
- ASASEC: Advisory System Against Sexual Exploitation of Children. European commission.
Financing entity: UE
Duration from: 16/10/2012 to: 16/10/2014
Main researcher: Enrique Alegre Gutiérrez
Quantity: 716.969,63 €
 - Age estimation from face to detect people with special needs assistance. Telefónica.
Financing entity: TELEFÓNICA
Duration from: 25/10/2013 to: 25/10/2014
Main researcher: Enrique Alegre Gutiérrez
 - Automatic assessment of fresh and criopreserved boar sperm through digital image segmentation, analysis and classification. Spanish Ministry of Science and Innovation.
Financing entity: MICINN
Duration from: 01/01/2010 to: 31/12/2012
Main researcher: Enrique Alegre Gutiérrez
Quantity: 84000 €
 - Proximal drops detection in boar spermatozoon tails through digital image processing. Microptic S.L.
Financing entity: MICROPTIC S.L.
Duration from: 01/10/2010 to: 31/08/2012
Main researcher: Enrique Alegre Gutierrez
Quantity: 10620 €
 - Application of Leading Technologies to Unmanned Aircraft Research and Development in Air Traffic Management (ATLANTIDA)
Main researcher: Jesús Gonzalo De Grado

Attended conferences

- XXXIV Jornadas de Automática, Terrassa, Spain, September 4-6 2013.
- International Conference on Similarity Search and Applications (SISAP), 2013.
- International Conference on Image Analysis and Recognition ICIAR. Aveiro, Portugal., 2012.
- Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA), Las Palmas de Gran Canaria, 2011.
- III Ecomas thematic conference on computational vision and medical image processing. Algarve, Portugal, 2011

Summer schools

- Summer schools International Computer Vision Summer School (ICVSS 2012). 3R's of Computer Vision: Recognition, Registration, Reconstruction. Sicily, Italy. 16-21 July 2012

- International Computer Vision Summer School (ICVSS 2011). Registration, Recognition and Reconstruction in Images and Video Sicily, Italy. 11-16 July 2011

Annex B

**SUMMARY OF THE THESIS IN
SPANISH**

**RESUMEN DE LA TESIS EN
CASTELLANO**

En cumplimiento del punto 3º del artículo 19 del Reglamento de las enseñanzas Oficiales de doctorado y del título de doctor de la Universidad de León, aprobado en Consejo de Gobierno el 25/9/2012, se adjunta un resumen en castellano de cada uno de los capítulos de esta tesis doctoral para que pueda admitirse a trámite.

1 Introducción

1.1 Motivación

En esta tesis se aborda el problema de monitorización del desgaste de la herramienta de corte utilizando nuevas técnicas de visión por computador basadas en descripción de forma y textura. Algunas ventajas de la incorporación de sistemas automáticos son, en primer lugar, que evitamos la subjetividad a la hora de tomar decisiones y en segundo lugar, las mejoras en términos de coste y de velocidad de procesamiento.

Reemplazar las herramientas en el momento adecuado es esencial debido al coste de las mismas y también a los costes indirectos derivados de la pérdida de tiempo al llevar a cabo el reemplazamiento. Por esta razón, la monitorización del estado de la herramienta se ha convertido en una pieza clave en la industria para optimizar las operaciones y aumentar la productividad. Otro de los objetivos de este trabajo es desarrollar propuestas que puedan ser implementadas utilizando dispositivos de bajo coste, fácilmente instalables y que precisen de un sencillo mantenimiento.

Este proyecto de investigación ha sido llevado a cabo con la colaboración de TECOI, una compañía de Castilla y León especializada en la construcción de máquina herramienta. Nuestra propuesta trata de mejorar estas máquinas mediante un método de predicción de la vida útil de la herramienta de corte que facilita el trabajo de los operarios y también la fiabilidad a la hora de hacer los cortes. Una de las principales ramas de producción es la creación de las estructuras de las turbinas de viento, por tanto, este proyecto tiene gran impacto en el área de prioridad de Castilla y León de medio ambiente y energía. Mejorando la producción de estas estructuras, podrían crear mejores turbinas y multiplicar la cantidad de energía generada por ellas.

Una de las últimas innovaciones de TECOI ha sido el diseño y fabricación de una máquina que resuelve el problema al que se someten las torres de los aerogeneradores. Las torres, debido al proceso de corte, el proceso de preparación de la soldadura,..., al estar creadas por grandes superficies curvadas, llegan a presentar microfisuras. TECOI ha eliminado este problema con su máquina TRF que corta con plasma o con oxyfuel dejando las soldaduras perfiladas porque va moliendo el

filo de las placas. Esta solución está compitiendo en el mercado mundial, pero aun presenta el problema de la posible rotura de las plaquitas de corte durante el proceso de cortado, por tanto, predecir la vida útil de la herramienta de corte antes de cada pasada, sirve de gran ayuda para mejorar esta máquina. Determinar el desgaste de la herramienta es una prioridad del proyecto DPI2012-36166 dentro del cuál está enmarcada esta tesis. El principal objetivo es producir piezas finales acordes con los requerimientos, procediendo a reemplazar la herramienta de corte por una nueva cuando llegue al final de su vida útil.

1.2 Objetivos

El principal objetivo es caracterizar el nivel de desgaste de plaquitas de corte en procesos de fresado analizando y clasificando imágenes digitales. Dado este objetivo general, definimos los siguientes objetivos particulares:

- Estudio de diferentes descriptores de imágenes digitales (forma, contorno, momentos, basados en textura,...) y su aplicación a conjuntos de datos ya existentes.
- Proposición de nuevos descriptores que aborden las debilidades detectadas en el análisis previo.
- Creación de conjuntos de imágenes con el fin de reproducir el escenario real de las máquinas.
- Categorización del estado de la herramienta usando técnicas de aprendizaje máquina y describiendo la región de desgaste de cada plaquita con los descriptores propuestos en este trabajo.
- Creación de un sistema automático y online, fácilmente adaptable a todo tipo de herramientas para predecir el desgaste de las plaquitas.

1.3 Contribuciones principales

Las principales contribuciones de esta tesis se pueden resumir como sigue:

- Se han creado tres nuevos conjuntos de datos para la monitorización del desgaste de la herramienta. El conjunto de datos de bordes de plaquitas, empleado para descriptores de momentos y contorno, que contiene información sobre la forma de la región de desgaste; el conjunto de datos de regiones de plaquitas, que distingue entre desgaste de los diferentes bordes de la plaquita, horizontales o verticales, y por último, el conjunto de datos de plaquitas de alta resolución, desarrollado para el análisis basado en textura.

-
- Se ha introducido un nuevo método llamado CPDH36R basado en el ya conocido CPDH y se ha probado en dataset clásicos de forma: MPEG y Kimia. La propuesta no solo obtiene mejores resultados en cuanto a precisión sino que también mejora al CPDH en términos de coste computacional.
 - Se ha propuesto un nuevo método llamado aZIBO para la recuperación de formas. Está basado en momentos y en características invariantes a la rotación. Se tomó como referencia el método ya existente ZMEG pero aZIBO añade invarianza a la rotación, lo cual lo hace mucho más robusto. El descriptor aZIBO fue evaluado en el problema del desgaste de las plaquitas obteniendo resultados que mejoran el método de base.
 - Se desarrolló una evolución más compleja de aZIBO, llamada B-ORCHIZ, para la evaluación del desgaste de las plaquitas de corte. El método añade nueva información relacionada con las orientaciones del borde de la forma de la región de desgaste lo que permite una mejor descripción y no solo es útil para la tarea específica de categorización del desgaste, sino también para conjuntos de datos de forma genéricos.
 - Se ha llevado a cabo un análisis de agrupamientos basado en dendrogramas para encontrar la mejor manera de clasificar las plaquitas de corte en función de su nivel de desgaste. La evaluación fue realizada usando B-ORCHIZ, aZIBO y ZMEG y B-ORCHIZ demostró ser el algoritmo con mayor poder de discriminación a la hora de describir las plaquitas.
 - Se ha llevado a cabo un estudio de la forma representativa de cada clase de desgaste para ayudar a los expertos a identificar más fácilmente los diferentes niveles de desgaste comparando las plaquitas nuevas a analizar con las formas representativas.
 - Se introduce un nuevo método basado en características de regiones. Se llama ShapeFeat y es un descriptor con muy bajo coste computacional que además requiere muy poca memoria para ser almacenado. El método propuesto alcanza muy buenos resultados en cuanto a la clasificación del desgaste utilizando el conjunto de datos de regiones de plaquitas.
 - Se han usado diferentes métodos de clasificación que fusionan descriptores. Los hemos empleado para poder tener en cuenta tanto la información de la forma como la información del contorno a la hora de clasificar el desgaste. También demostramos el buen funcionamiento de ShapeFeat y B-ORCHIZ cuando se combinan.
 - Se propone un nuevo sistema online para la evaluación automática del desgaste de las plaquitas en procesos de fresado. El sistema es de muy bajo

coste, completamente automático y fácilmente adaptable a cualquier tipo de máquina de corte. Se basa en la división de las imágenes de las plaquitas en pequeñas regiones para describir cada una independientemente. El método de clasificación tiene en cuenta la proporción de regiones desgastadas y de esta manera puede determinar la clase de la plaquita entera.

1.4 Organización del resto del documento

En esta sección se describe la estructura de este resumen de la tesis doctoral. Este primer capítulo introductorio está centrado en motivar el trabajo presentado en esta disertación, los objetivos principales y las contribuciones originales. El resto de los capítulos están organizados de la siguiente manera:

El capítulo 2 contiene una revisión de los métodos de descripción de forma así como de los métodos basados en monitorización del estado de la herramienta, y de los métodos basados en sistemas de visión.

El capítulo 3 presenta los conjuntos de imágenes de forma con los que han sido probados nuestros métodos y también la creación de tres conjuntos de imágenes para la aplicación real de monitorización del desgaste de la herramienta: conjunto de imágenes del filo de las plaquitas, conjunto de imágenes de regiones y conjunto de imágenes a alta resolución para su descripción mediante textura.

Los dos nuevos descriptores de forma basados en contorno y momentos se describen en el capítulo 4. Estos métodos fueron probados en los conjuntos de imágenes de forma genéricos, demostrando su efectividad.

El capítulo 5 explica dos nuevas propuestas de métodos de descripción: aZIBO y B-ORCHIZ, que combinan las ventajas de los descriptores globales y locales. Fueron empleados en la tarea de monitorización del desgaste consiguiendo resultados prometedores.

En el capítulo 6 se presenta el descriptor ShapeFeat, para caracterizar el desgaste de la herramienta de corte de procesos de fresado. También se evaluaron métodos de clasificación para combinar nuestra nueva propuesta ShapeFeat con B-ORCHIZ, obteniendo aun mejores resultados que usando cada método por separado.

El método nuevo de descripción basado en la división de la imagen en patches se detalla en el capítulo 7.

Finalmente, el capítulo 8 contiene un resumen con las conclusiones de esta tesis y da una visión sobre las posibles líneas de trabajo futuras.

2 Revisión del estado de la técnica

2.1 Monitorización del desgaste de la herramienta

En esta tesis abordamos la tarea de monitorizar el desgaste de las herramientas en procesos de fresado utilizando técnicas de visión por computador y aprendizaje máquina. Los sistemas de monitorización de la condición de las herramientas son útiles para la detección prematura de fallos en el sistema y comprobar la estabilidad del proceso de cortado, motivos por los cuales la tolerancia de la máquina se mantiene hasta los límites aceptables (Dimla and Dimla, 2000). Es importante remarcar que no solo el coste de las herramientas es significativo, sino también los costes indirectos derivados del tiempo improductivo necesario para llevar a cabo los reemplazamientos. Optimizar estas operaciones de reemplazamiento producirá una mejora significativa en cuanto a la eficiencia y competitividad.

Los métodos más conservadores normalmente reemplazan las herramientas antes de que sea realmente necesario para evitar el alto coste de que fallen, de manera que se retiran cuando aún les queda algo de vida útil. En (Malekian et al., 2009) se estima que los fallos en las herramientas típicamente representan el 20 % del tiempo que una máquina está inactiva. Sin embargo, reemplazar las herramientas en el momento óptimo es esencial por los costes de las herramientas y los costes indirectos que conlleva su cambio. Por esta razón, la monitorización del desgaste de la herramienta se ha convertido en una parte muy importante para la industria para optimizar las operaciones y así aumentar la productividad.

Los sistemas existentes de monitorización del desgaste de la herramienta se pueden resumir en dos grandes categorías: métodos indirectos y métodos directos.

2.1.1 Métodos indirectos

Este tipo de métodos estiman el desgaste midiendo variables como la temperatura de la herramienta (Wang, Ming and Chen, 2016), las fuerzas de corte (Wang, Yang, Xie and Zhang, 2014; Azmi, 2015; Kaya et al., 2011), vibraciones (Li, 2002; Rao et al., 2014) o emisiones acústicas (Scheffer and Heyns, 2001), y también combinaciones de métodos como medidas ultrasónicas y ópticas, dimensiones de las piezas, análisis del estrés,... Todos estos métodos están detallados en (Dimla and Dimla, 2000).

En (Bhuiyan et al., 2016), se usa un sensor de emisiones acústicas para computar características y predecir así el desgaste de la herramienta. También en (Zhou et al., 2011), utilizan emisiones acústicas en sensores embebidos y en (da Silva et al., 2016), se introdujo un nuevo parámetro basado en la energía de emisión de señales (entre 100 y 300 kHz) para mejorar la respuesta.

La temperatura de la herramienta es otro factor a tener en cuenta ya que los

procesos de corte de metales generan gran cantidad de calor. En (Ceau et al., 2010) se lleva a cabo un estudio experimental sobre las medidas de temperatura en procesos de corte y más recientemente, en (Chryssolouris et al., 2012), se describe una metodología para medir la temperatura usando un pirómetro bi-color de fibra óptica. Utilizaron dos posiciones diferentes de la fibra para determinar la temperatura en dos localizaciones diferentes, el centro y la esquina exterior del taladro. Desafortunadamente, los modelos existentes están establecidos para predecir el desgaste en procesos continuos de torneado y no son adecuados para procesos interrumpidos de fresado. Para solucionar este problema, en (Wang, Ming and Chen, 2016), los autores primero determinan un umbral de temperatura y después proponen un modelo matemático para predecir el desgaste a través de fuerzas y temperaturas.

También está muy establecido el uso de fuerzas de corte a la hora de determinar el estado de las herramientas (Wang, Yang, Xie and Zhang, 2014; Azmi, 2015; Kaya et al., 2011). Además de los métodos ya citados, en (Nouri et al., 2015), se desarrolla un nuevo método para detectar el desgaste en tiempo real midiendo los coeficientes durante el proceso de corte. Está comprobado que el comportamiento de estos coeficientes es independiente de las condiciones del proceso de corte.

En cuando a las vibraciones, se producen por variaciones cíclicas en los componentes dinámicos de las fuerzas de corte. Un ejemplo de trabajo en el que utilizan este tipo de señales es en (Bovic Kilundu, 2011). Se lleva a cabo un análisis del espectro, se acopla a un filtro de paso bajo para permitir la definición y la extracción de características sensibles al desgaste y estas características son posteriormente definidas, en diferentes bandas de frecuencia, por los coeficientes de Fourier. También en (Prasad and Babu, 2017) se lleva a cabo una correlación entre la amplitud de las vibraciones y el desgaste de la herramienta.

También surgen trabajos en los que se combinan diferentes tipos de estos métodos. En (Dimla, 1999) se elabora una investigación en la que se combinan medidas de fuerza con señales de vibraciones. En (Segreto et al., 2013) se emplean múltiples sensores que comprenden mediciones de fuerzas de corte, emisiones acústicas, vibraciones, etc. La extracción de características la realizan procesando las señales detectadas por los sensores para reducir su dimensionalidad y posteriormente utilizan una red neuronal para la clasificación. Otro ejemplo donde se combinan fuerzas de corte con emisiones acústicas es en (Choi et al., 1999). Sin embargo, todos estos métodos basados en señales y sensores no dan la suficiente información para llevar a cabo un reemplazamiento óptimo de las herramientas de corte porque la relación entre el desgaste y las variables observadas depende de las condiciones de corte y, en general, no se suelen conocer. Aunque estos métodos son los más populares, la precisión que se alcanza con ellos está muy afectada por el ruido que contengan las señales en entornos industriales (Kassim et al., 2007).

2.1.2 Métodos directos

En cambio, los métodos directos, generalmente basados en sistemas de visión artificial, tienen la ventaja de que miden cambios reales de geometría en la herramienta, ofreciendo una mayor precisión (Loizou et al., 2015). En (Kurada and Bradley, 1997) y (Dutta et al., 2013) se presenta un estudio de métodos de visión artificial para la monitorización de desgaste. Algunos de estos métodos se basan en la forma del contorno, otros en la textura, en propiedades de la forma, o en combinaciones de dos o más de estas técnicas.

En las siguientes dos secciones se hará un breve resumen de los principales métodos de descripción de forma y textura para iniciar al lector en la temática y ayudar a la comprensión de las contribuciones de esta disertación.

2.2 Descripción basada en forma

En las últimas décadas, el rápido incremento de los dispositivos de captura de imágenes ha hecho que el número de ellas crezca exponencialmente. Y al mismo tiempo, crece también la demanda de indexar, almacenar y recuperar estas imágenes digitales. Se han desarrollado numerosos métodos de descripción de forma. Las imágenes generalmente se suelen describir por características basadas en su contorno, color, forma o textura. Dos de los factores clave en un sistema de reconocimiento de imágenes son su eficiencia y su coste computacional. La recuperación de imágenes a través de la forma de la misma ha ido ganando importancia en la comunidad de reconocimiento de patrones y procesamiento digital de imágenes centrándose y mejorando estos dos factores clave. Algunos trabajos se han centrado en la utilización de medidas de distancia entre formas. Por ejemplo, Nasreddine et al. definieron una distancia multiescala basada en geodésicos en el espacio de la forma, invariante a traslaciones, rotaciones y escalado (Nasreddine et al., 2010). Ćurić et al. (Curic et al., 2014) propusieron otra medida de distancia donde cada punto de cada conjunto, se pondera dependiendo de su distancia al complementario del conjunto.

Hay bastantes tipos de métodos para la representación de las formas dependiendo de la información extraída para construir el descriptor. En el estudio realizado por Yang et al. (Mingqiang et al., 2008) se presentan diferentes tipos de representación de formas: métodos basados en parámetros básicos, métodos de funciones uni-dimensionales, aproximaciones poligonales, características de integración espaciales, momentos, y finalmente dominios de transformación de la forma.

2.2.1 Parámetros básicos de forma

Generalmente solo se pueden distinguir formas que son muy diferentes entre sí. Sin embargo, hay múltiples campos en los que este tipo de características son muy útiles y además también resultan útiles combinándolas con otro tipo de características.

El radio de circularidad es una característica que representa cómo es de similar una forma a un círculo. En (Wang, Chang, Zhou, Li, Liu and Guo, 2016), este radio se usa para identificar cinco tipos diferentes de glóbulos blancos evaluando su forma en imágenes captadas mediante un microscopio. Este radio se puede definir como el ratio del área de la forma entre el área del círculo que tiene el mismo perímetro.

Otra característica de forma básica ampliamente usada en la literatura es la convexidad. Se define como el ratio del perímetro de la región convexa más pequeña que incluye a la figura (convex hull) entre el contorno original.

En el trabajo desarrollado Yasseen et al. (Yasseen et al., 2016), propusieron un método para la recuperación de formas basado en dividir la imagen en partes visuales y describirlas usando parámetros de forma como el área, circularidad, convexidad,... Recientemente, Fishbaugh et al. (Fishbaugh et al., 2017) propusieron un modelo que desacopla los parámetros de deformación de las representaciones de forma específicas, permitiendo que la complejidad del modelo refleje la naturaleza de los cambios de forma, en lugar de muestrear los datos.

2.2.2 Métodos de funciones uni-dimensionales

Este tipo de métodos captura las características perceptuales de la forma. Aunque se pueden utilizar solos para describir formas, normalmente se usan en combinación con otros métodos de extracción de características como por ejemplo Fourier o Wavelet.

En (Condorovici et al., 2015), los autores imitan la percepción humana a bajo y medio nivel. En su trabajo, usan la energía del filtro de Gabor para la descripción y además proponen, también para la descripción, el uso del volumen del color dominante (DCV) y los marcos extraídos usando la teoría del anclaje. Para la clasificación, se realizó una combinación de Perceptron de Multi-Capa (MLP) con Máquinas de Vector de Soporte (SVM).

La curvatura es una característica muy representativa para el ojo humano a la hora de diferenciar formas. También contiene información perceptual y se ha probado que es muy útil para el reconocimiento de formas. En (Czerniawski et al., 2016), los autores presentan un método automatizado para localizar y extraer carretes de tubos en exploraciones de nubes de puntos desordenadas. Se basa en la estimación de la curvatura de los datos a nivel local y en hacer agrupamientos. Finalmente, la correspondencia se hace mediante BoF (Bag of Features).

La mayoría de estos descriptores tienen problemas relacionados con el coste computacional. Los descriptores de Fourier son considerados como un método de bajo coste computacional y basándose en esta información, Elghazal et al. (El-ghazal et al., 2012), propuso un nuevo descriptor de Fourier basado en curvaturas (CBFD) para la recuperación de formas. Este método es capaz de capturar los detalles de la curvatura de la forma y mejorar la eficiencia de la recuperación en el proceso de correspondencia de las formas.

Otro método interesante y ampliamente usado en la representación de formas es la función del área. Cuando los puntos del contorno van cambiando a lo largo del mismo, el área del triángulo formado por dos puntos de contorno consecutivos y su centroide, también cambian. Con esto, se puede crear una función del área y usarla como una representación de la forma.

Todos estos métodos basados en funciones uni-dimensionales derivan del contorno de las formas. Para obtener invarianza a la traslación, generalmente se definen con valores relativos. La invarianza a la escala es más fácil y se obtiene simplemente normalizando el descriptor final. Los cambios en la orientación normalmente se evitan en el proceso de emparejamiento (matching) desplazando los valores del descriptor hasta obtener la cadena de mínima distancia. De todas formas, son bastante sensibles al ruido y pequeños cambios en los puntos del contorno pueden causar grandes errores en la descripción. Por este motivo, estos métodos se suelen usar en combinación con otros.

2.2.3 Aproximacion poligonal

Estos métodos se usan para disminuir la variación sobre el borde de la forma, intentando capturar toda la información de la misma. En general hay dos tipos diferentes: mezclar y dividir. Los métodos de mezclar van añadiendo sucesivamente píxeles a un segmento si el nuevo píxel a añadir no causa que el segmento se desvíe demasiado de una línea recta. El umbral se fija eligiendo dos puntos del contorno y calculando el error para cada punto del contorno entre esos dos puntos seleccionados. Si el error excede un umbral, elegimos dos puntos más cercanos para crear el segmento y comenzamos de nuevo el proceso.

Por otra parte, los métodos de división, comienzan dibujando una línea desde un punto del contorno hasta otro. Después, se computa la distancia perpendicular desde cada punto de contorno y si excede un umbral, se divide esa línea inicial en dos segmentos por el punto de máxima distancia.

En (Madrid-Cuevas et al., 2016), se combinan las dos estrategias (mezclar y dividir). El algoritmo que proponen obtiene aproximaciones poligonales de los contornos y no tiene que ajustar ningún parámetro, lo cual nos da el mejor balance entre fidelidad y eficiencia y tiene una complejidad relativamente baja. El método se basa en el análisis de la concavidad y convexidad del árbol del contorno y después

utiliza una estrategia eficiente de división/mezcla.

En los últimos años se han desarrollado bastantes métodos basados en puntos salientes (Kim and Park, 2014) y curvatura de la forma para describir imágenes.

En (Andaló et al., 2010) se usa IFT pero los autores también introducen el concepto de escala de tensor, que es un parámetro morfométrico que unifica la representación de la estructura, la orientación y la anisotropía. (IFT es una generalización del algoritmo de Dijkstra: es un enfoque basado en grafos que devuelve un mapa raíz, un mapa de costes y un mapa de etiquetas que contiene la información relevante de los puntos de contorno y su relación con sus puntos de área de influencia).

También se han usado técnicas de puntos salientes en (Pedrosa et al., 2013). La técnica presentada en este trabajo consta de un detector de puntos salientes, una representación que usa la posición angular relativa y también un valor de curvatura. La técnica que presentan es robusta frente al ruido y además obtiene buenos resultados diferenciando entre formas que pertenecen a clases diferentes pero que son muy similares entre ellas.

Recientemente, Laiche et al. (Laiche et al., 2014), propusieron un método para descripción y reconocimiento de formas bidimensionales. Primero emplean aproximaciones poligonales para representar la forma subyacente y después usan el modelo de mínimos cuadrados. Cada parte de la aproximación es asociada con una curva polinomial de grado dos. Las curvas que se obtienen se normalizan y son invariantes al escalado, rotación y traslación. Finalmente, basándose en la similitud de las formas resultantes, se define un modelo de similitud entre la forma a evaluar y el modelo de referencia. Los autores también propusieron un algoritmo de correspondencia de dos pasos.

2.2.4 Momentos

En el ámbito del reconocimiento de imágenes, los momentos no se utilizan como en matemáticas o en física. Se emplean para el análisis de objetos. Los momentos de Zernike son ampliamente usados en este ámbito de recuperación de imágenes y los momentos complejos de Zernike se derivan de los polinomios de Zernike ortogonales. Tienen las siguientes ventajas:

- Invarianza a la rotación.
- Son robustos al ruido y a pequeñas variaciones de la forma.
- Baja redundancia.

Así mismo también tienen desventajas como que la imagen tiene que ser transformada al dominio en el que está definido el polinomio ortogonal, las

integrales tienen que ser transformadas en sumas discretas aproximadas y generalmente tienen alto coste computacional.

Los momentos de Zernike siguen siendo ampliamente usados en la actualidad.

En 2016, (Wu et al., 2016), llevó a cabo un trabajo muy interesante para detectar las acciones de estudiantes mediante una cámara. El método está basado en la descripción, mediante momentos de Zernike, de imágenes que representan tres estados diferentes: con las manos en alto, de pie y sentado. Además, esta técnica fue combinada con el flujo óptico Lucas-kanade para describir la dirección y velocidad de los movimientos. Los resultados demostraron eficacia y robusted identificando en los estudiantes estas tres acciones.

A principios de 2017, Kumar et al. propusieron un método para identificar plantas basándose en imágenes de sus hojas. Para ello combinaron descriptores de forma y de textura. Como descriptor de forma se usó un momento ortogonal (Krawtchouk moments) y como descriptores de textura, HOG y Gabor. El sistema se entrenó con máquinas de vector soporte (Pradeep Kumar et al., 2017).

Recientemente, (Bolourchi et al., 2017), describió imágenes usando los momentos radiales de Chebyshev fusionados con los momentos de Zernike. Los resultados experimentales que la combinación de ambos momentos mejoraba la clasificación en más de un 6%.

Todos estos estudios nos muestran que los momentos son, por lo general, concisos, robustos y fáciles de calcular. Además la mayoría de ellos son invariantes a la rotación, escalado y traslación. Sin embargo, debido a su naturaleza global, tienen problemas cuando se trata de correlacionar momentos de gran orden con características salientes de las formas a evaluar.

2.2.5 Dominios de transformación de la forma

Incluyen métodos que tratan de transformar los objetos detectados o la imagen entera. El método más eficiente dentro de este grupo es la transformada de Fourier.

En (Sokic and Konjicija, 2016), se propone un nuevo método para extraer los descriptores de Fourier que preserven la fase de los coeficientes. Introducen unos puntos específicos llamados “pseudomirror points” y los usan como referencia de orientación de la forma. En esta misma línea podemos encontrar más trabajos recientes, que todavía utilizan Fourier con resultados satisfactorios (Shu et al., 2015; Ai et al., 2013).

Otro método importante es la transformada de Wavelet, que descompone una curva en componentes a diferentes escalas de manera que las escalas más grandes contienen aproximaciones de la forma más globales mientras que las escalas más pequeñas se encargan de los detalles. Esta técnica tiene propiedades interesantes como como invarianza al escalado, traslaciones y rotaciones, representación multi-resolución, estabilidad y correlación espacial. Aunque sea un método antiguo,

al igual que Fourier, todavía se usa obteniendo buenos resultados (Shahsavari et al., 2017; Masoumi et al., 2016; Hadizadeh, 2015).

Algunos grupos de investigación han mezclado las transformadas Fourier y Wavelet con el objetivo de desarrollar métodos más complejos y eficientes (Yuan et al., 2014).

Otros métodos recientes que tratan este problema de reconocimiento de formas se basan en medidas de similitud u otros métodos de clasificación. Biswas et al. presentaron un método de indexación de formas para grandes conjuntos de imágenes que mejoraba el resto de métodos del estado del arte (Biswas et al., 2010). Este método es invariante a articulaciones, transformaciones rígidas,...y está basado en relaciones geométricas entre pares de puntos de interés de la forma. Shu y Wu (Shu and Wu, 2011a) introdujeron un descriptor de contorno basado en la distribución de los puntos del objeto en coordenadas polares. Tiene un coste computacional bajo y es invariante a escalados, sin embargo, solo puede tratar con imágenes con el contorno cerrado y no es invariante a rotaciones. Más adelante, otro método de recuperación de formas, pero que sí que es invariante a rotaciones y escalado, fue el presentado en (Mukanova et al., 2014).

En los últimos años se ha hecho bastante frecuente la combinación de descriptores globales con locales. En (Singh and Pooja, 2012) se utiliza la transformada radial angular (ART) como descriptor global y la transformada de Hough como descriptor local. En (Anuar et al., 2013) también se usa una combinación de descriptor global y local pero en este caso, Zernike funcionaría como descriptor global, y EGCM (Matriz de coocurrencia de gradientes) como descriptor local.

2.2.6 Monitorización del desgaste de la herramienta mediante descriptores de forma

Existen muy pocos métodos basados en descripción de forma para este problema y además, los métodos de contorno existentes no dan la información suficiente como para determinar el nivel de desgaste porque hay muchos otros factores además de la forma, que lo caracterizan.

Recientemente, la comunidad científica ha desarrollado nuevos métodos que usan la información de la forma para describir las imágenes (Shu and Wu, 2011b; Laiche et al., 2014). Estas características de forma se extraen de las imágenes y después se procesan con modelos de clasificación o regresión para obtener información relevante sobre el estado de la herramienta. En (Dominguez-Caballero et al., 2016) se presenta un método para describir el desgaste de herramientas de cerámica. Los métodos de descripción de forma también han sido ampliamente usados en el ámbito de la monitorización del desgaste para asegurar el reemplazamiento óptimo de las herramientas en procesos de torneado. En esta línea, Barreiro et al. propone el uso de los momentos de Hu y Legendre (Barreiro

et al., 2008). Nuestro problema es diferente al detallado en (Barreiro et al., 2008) en el sentido de que nosotros determinamos el desgaste en procesos de fresado y no en procesos de torneado. En fresado, el desgaste de la forma se distribuye uniformemente a lo largo del filo de la plaquita mientras que en los procesos de torneado el desgaste se muestra de forma cóncava. En este sentido, medir el desgaste de la herramienta en procesos de fresado en el filo de corte, es una tarea más complicada debido a que la distribución del desgaste no es la típica y además la geometría del cabezal es más compleja.

2.3 Descripción basada en textura

Hoy en día, si nos centramos en el procesamiento de imágenes digitales, los métodos de descripción de textura son los más empleados. Las características de textura, generalmente se calculan desde la distribución estadística de las intensidades de los píxeles con respecto a la posición relativa de cada uno de ellos. En las siguientes secciones vamos a evaluar diferentes maneras de extraer estas características y su aplicabilidad en los procesos de monitorización del desgaste de las herramientas de corte.

2.3.1 Matrices de coocurrencia de niveles de gris

El método de la matriz de coocurrencia de niveles de gris (GLCM) es una manera de extraer información útil de segundo orden estadístico. GLCM es una matriz donde el número de filas y de columnas es igual al número de niveles de gris de la imagen (generalmente 256). El elemento $M(i,j)$ de la matriz es la frecuencia relativa con la que dos píxeles separados por una distancia fijada y con una orientación determinada, aparecen en un vecindario, uno con intensidad i y el otro con intensidad j .

A lo largo de los años se han desarrollado muchos trabajos que han utilizado este método de extraer información para construir sus descriptores de características.

En 1973, Haralick y su equipo de investigación desarrollaron un método (Haralick et al., 1973) basado en características extraídas de GLCM y evaluado con tres conjuntos de imágenes de areniscas, fotografías aéreas e imágenes ERTS (Earth Resource Technology Satellite). En este trabajo extraen 14 características diferentes de la matriz de coocurrencia de cada imagen: momentos de velocidad angular, contraste, correlación, suma de cuadrados, momentos de diferencias de cuadrados, momentos de diferencias inversas, suma media, suma de varianzas, suma de entropías, entropía, diferencia de varianza, diferencia de entropía, información de medidas de correlación y coeficiente de correlación máxima.

Los resultados demuestran una alta eficiencia y hoy en día las características de Haralick extraídas de GLCM todavía son ampliamente usadas para descripción

de textura. In (Zayed and Elnemr, 2015), Zayed y Elnemr propusieron un nuevo sistema de evaluación basado en estas características de Haralick para discriminar entre diferentes tipos de anomalías en el pulmón.

En 1980, Conners and Harlow demostraron que GLCM es una técnica más poderosa que GLDM (Gray Level Difference Matrix), GLRLM (Gray Level Run Length Method) y que PSM (Power Spectral Method) (Conners and Harlow, 1980).

Más recientemente, en (Mohanaiah et al., 2013), Mohanaiah y su equipo propusieron una nueva técnica de extracción de características GLCM basada en la evaluación de solo cuatro características, lo cual mejora no solo la velocidad de procesamiento sino también la precisión de la clasificación en imágenes de dibujos animados.

2.3.2 Patrones binarios locales (LBP)

LBP (Local Binary patterns) fue propuesto por Ojala y su grupo de investigación (Ojala and Pietikainen, 1996). Presentan un método basado en la extracción de un código binario para cada pixel teniendo en cuenta la intensidad de nivel de gris de sus vecinos, asignando el valor 1 si el nivel de intensidad es mayor que el pixel central del vecindario y asignando el valor 0 en caso contrario. Después de esto se crea un histograma para representar la frecuencia de las diferentes posibilidades de código binario que forman el descriptor LBP. En (Ojala and Pietikainen, 2002), los mismos autores propusieron dos extensiones del LBP original. Primero, lo definieron para diferentes configuraciones de vecinos y después definieron los patrones uniformes: un código binario LBP en el que cambian menos de tres valores. Por ejemplo, 00001111 o 11000001 son patrones uniformes mientras que 10010100 no lo es.

LBP ha sido ampliamente usado en la literatura debido a su bajo coste computacional y a su alto poder discriminativo.

En (Doost and Amirani, 2013), se propone un nuevo método que combina LBP con transformadas Wavelet continuas para clasificar texturas de un conjunto de imágenes de textiles formado por 64 imágenes repartidas en 12 clases diferentes.

En el trabajo llevado a cabo por Garcia-Olalla et al (García-Olalla et al., 2015), se desarrolla un nuevo método para determinar la integridad de acrosomas de espermatozoides de berraco combinando LBP con otros métodos de descripción de textura y forma y se alcanzan tasas de acierto cercanas al 100 %.

En la sección de metodología se va a profundizar en el descriptor LBP y algunas de sus variantes más importantes.

2.3.3 Monitorización del desgaste de la herramienta mediante métodos de descripción de textura

En (Datta et al., 2013) se propone un método basado en el análisis de texturas y en la teselación de Voronoi.

En (Dutta et al., 2012), se lleva a cabo una adquisición online de imágenes de la superficie. Después estas imágenes se analizan usando una matriz de co-ocurrencia (en escala de grises). A partir de este trabajo los autores propusieron el uso de la transformada discreta de Wavelet para describir superficies torneadas (Dutta et al., 2016b).

En (Li and An, 2016), se propone otro método nuevo también basado en la matriz de co-ocurrencia de niveles de gris (GLCM). Los autores la utilizan para extraer la mejor ventana de análisis de textura de una plaquita, mostrando que la entropía permanece monótona a medida que aumenta el desgaste de la herramienta, lo que demuestra que la entropía se puede utilizar eficazmente para controlar ese desgaste. También se utiliza GLCM en (Sun et al., 2017).

En (García-Olalla et al., 2017), se extraen descriptores basados en LBP, del conjunto de imágenes Outex, y se combinan con LOSIB (Local Oriented Statistical Information Booster) pero en este caso en concreto, los resultados que se obtienen con la combinación no son los esperados debido, probablemente, a la naturaleza del conjunto de datos, creado para probar métodos invariantes a la rotación.

Recientemente, en (Antić et al., 2018), se utiliza el módulo del espectro de la transformada Discreta de Fourier de Corto Plazo (STDFT). Se obtiene de la emisión de señal de los sensores de vibración. Esto se hace identificando la escala de tiempo de STDFT como la primera dimensión, y la escala de frecuencia como la segunda dimensión de la imagen texturizada en particular. No utilizan un sistema de fotografía, sino que tienen en cuenta la potencia de la descripción de la textura aplicada sobre las fuerzas de vibración.

3 Datasets

En esta sección se presentan y describen los conjuntos de imágenes tanto de formas como los de plaquitas, creados por nosotros, que se citarán a lo largo de esta tesis y que fueron usados para probar todas nuestras propuestas.

3.1 Datasets de recuperación de forma

Para evaluar nuestros descriptores se usaron tres datasets clásicos ya existentes. Kimia25, Kimia99 y un subconjunto del conocido MPEG7 compuesto por 400 imágenes, llamado MPEG2. Kimia25 consta de seis clases. Cada una de ellas contiene cuatro formas excepto la clase de las manos, que contiene cinco muestras.

Kimia99 tiene nueve clases, cada una de ellas compuesta por once formas. Los conjuntos de imágenes de Kimia25 y Kimia99 se pueden ver en la Figura 1.

MPEG7 contiene setenta clases con veinte muestras por clase. MPEG2 es un subconjunto de MPEG7 compuesto por veinte de las clases de MPEG7. En la Figura 2 se pueden ver algunos de los ejemplos de las imágenes que las componen.

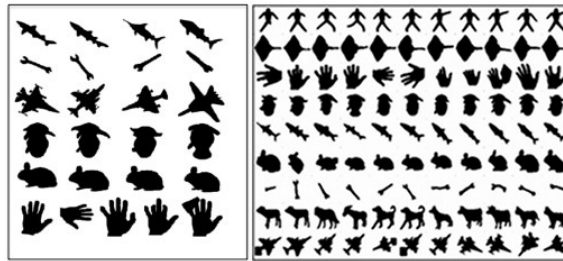


Figura 1: a) Conjunto de imágenes Kimia25 compuesto por 25 muestras repartidas en 6 clases. b) Conjunto de imágenes Kimia99 compuesto por 99 muestras repartidas en 9 clases.

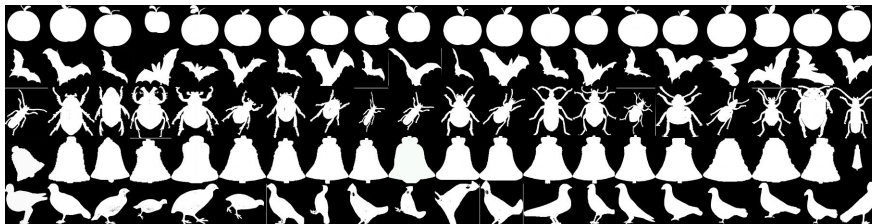


Figura 2: Ejemplo de 4 clases con 20 muestras del conjunto de imágenes MPEG2.

3.2 Conjunto de imágenes de filo de plaquitas

El objetivo de esta tesis es caracterizar el estado de las plaquitas de corte basándonos en la forma de la región de desgaste. Para llevarlo a cabo, hemos adquirido un conjunto de imágenes de plaquitas procesando imágenes de 53 herramientas como las que se muestran en la Figura 3.

Usando las imágenes en escala de grises en vez de imágenes en 3D (como hacen otros autores), podemos medir la región de desgaste en tiempo real, lo que hace posible desarrollar nuestra propuesta en un entorno real.

Las imágenes fueron tomadas con una cámara monocromática, modelo Genie M1280 1/3" con una óptica AZURE de 25mm. El enfoque y la apertura son manuales. El sensor tiene una resolución de 1280 × 960 píxeles.

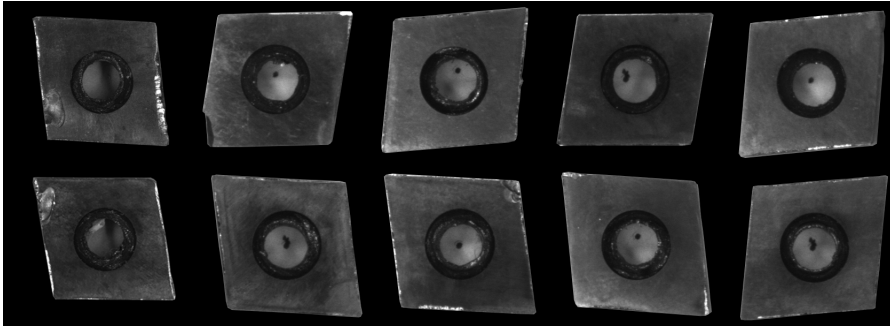


Figura 3: Ejemplo del conjunto de imágenes. Las plaquitas de la izquierda presentan alto desgaste mientras que las de la derecha presentan bajo desgaste teniendo siempre en cuenta el filo izquierdo de corte.

Para mejorar el contraste de la imagen en blanco y negro se utilizaron dos barras de iluminación LED roja (BDBL-R8216H). Las imágenes fueron tomadas con las plaquitas desmontadas del cabezal de corte, posicionadas sobre un fondo negro uniforme.

Cada plaquita, capturada como una imagen en escala de grises y con el fondo enmascarado, fue sometida a una etapa de preprocesamiento que dio lugar a cuatro imágenes, una por cada filo de corte. Por tanto, recortamos cada uno de los fillos de corte y rotamos los dos que son verticales, de manera que obtuvimos cuatro imágenes, cada una con un filo de corte en posición horizontal. A continuación se explica el proceso de preprocesamiento:

En primer lugar se eliminó la porción central de la plaquita, por tanto, necesitábamos obtener el centro y la longitud del radio de la porción. Para determinar el centro del círculo se binarizó la imagen usando un umbral de 0.01 obtenido empíricamente de manera que los píxeles de la plaquita valiesen 1 y los del fondo, 0. El centroide de este área es considerado como el centro del círculo central. El radio del círculo será $R_c = D/4,92$ donde D es la longitud de la diagonal mayor de la plaquita y el valor del ratio 4,92 se obtiene a partir del tamaño del círculo central con relación a la plaquita, que es el mismo en todo el conjunto.

Después de eso se extraen cuatro regiones con un ancho de $1/6$ de la plaquita y la misma altura que contienen los cuatro fillos de corte correspondientes. A continuación se explica cómo se obtuvo el filo izquierdo de corte. El resto se obtiene de la misma manera pero rotando la imagen original 90° , 180° o 270° , según corresponda.

Usando la imagen en escala de grises resultante del paso anterior, aplicamos un filtro Sobel vertical con kernel de 3×3 para detectar los contornos. Estos contornos fueron dilatados y sometidos a operaciones de apertura usando un

elemento estructural de forma cuadrada y tamaño 3 y posteriormente proyectados verticalmente en el eje horizontal. El primer elemento no nulo en esta proyección indica la coordenada x donde empieza el filo de corte. Después, la imagen se recortó desde esta posición x con una anchura de 100 píxeles y la misma altura que la imagen original. Se añadió un margen paramétrico de 25 píxeles a cada lado para incrementar el área recortada porque las pruebas experimentales demostraron que algunas plaquitas perdían algunos píxeles pertenecientes al filo de corte en el proceso de recorte de imagen.

Como las plaquitas tienen forma romboide y sus fillos de corte no están alineados con los ejes horizontal y vertical, estos también se rotaron. Para llevar a cabo la rotación se aplicó un filtro Sobel horizontal (kernel 3×3) seguido de una dilatación utilizando el mismo elemento estructural que en la etapa anterior. La imagen resultante se filtró para eliminar los objetos pequeños resultando en una imagen binaria que contiene el filo de corte. Finalmente, se estimó la orientación del eje mayor de la elipse que contiene al filo y el eje es rotado compensando esta orientación y dejándolo en posición horizontal.

Una vez que se han obtenido las regiones con los fillos de corte, el siguiente paso consiste en extraer el desgaste de cada imagen. En nuestro caso es más importante una buena segmentación del desgaste que la velocidad que nos daría un proceso automático pero menos fiable, por tanto, el ground truth se hizo de manera manual. Al final de todo este proceso se obtiene un conjunto compuesto por 212 imágenes. Se puede ver un ejemplo en la Figura 4

Figura 4: Ejemplo del conjunto de imágenes del filo de corte.

3.3 Conjunto de imágenes de regiones de plaquitas

El conjunto de imágenes de regiones se creó siguiendo el mismo proceso descrito en la sección previa. Las imágenes recortadas normalmente están formadas por dos tipos diferentes de regiones de desgaste: regiones completas y regiones incompletas. Las completas corresponden con el filo de corte horizontal mientras que las incompletas representan a los fillos que convergen en la horizontal. Para determinar el desgaste del filo completo por sí mismo y también cómo está influenciado por los fillos incompletos, se dividió el dataset completo en dos: Insert-C, formado por los fillos completos que se corresponden con el desgaste horizontal, e Insert-I, con los

filos incompletos, orientados verticalmente en la imagen. Se puede ver un ejemplo de este conjunto de imágenes en la Figura 5

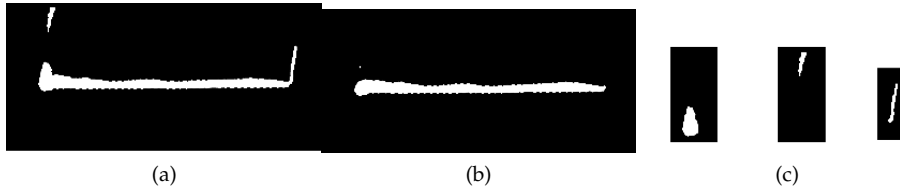


Figura 5: (a) Imagen que contiene los dos tipos de filos: Filo completo, posicionado en horizontal y filos incompletos en posición vertical. (b) Filo completo (c) Filos incompletos que representan los filos verticales en (a).

Un experto clasificó visualmente todas las regiones de desgaste del conjunto de imágenes para generar el ground truth, llevando a cabo dos tipos de calificaciones dependiendo de cuántas clases de desgaste se tuvieran en cuenta: Dos (Bajo desgaste (L) con 260 imágenes y alto desgaste (H) con 313 imágenes) o tres, (Bajo desgaste (L) con 126 imágenes, medio desgaste (M) con 260 imágenes y alto desgaste (H) con 187 imágenes).

3.4 Conjunto de imágenes en alta resolución para la descripción de texturas

No existía ningún conjunto de imágenes disponible públicamente con la suficiente calidad y cantidad de imágenes como para poder llevar a cabo una descripción basada en el análisis de texturas. Por esta razón, creamos un nuevo conjunto compuesto de 254 imágenes procedentes de un cabezal de corte de fresado. El cabezal tiene forma cilíndrica y contiene 30 plaquitas agrupadas de 5 en 5 diagonalmente. Hemos usado una cámara monocromática Genie M1280 1/3" con una resolución activa de 2592×1944 , lentes AZURE-2514MM con una distancia focal de $25mm$. La cámara se controla mediante una Raspberry Pi que puede ser fácilmente integrada en cualquier sistema de monitorización de desgaste debido a su tamaño pequeño. Para conseguir unas condiciones de iluminación óptimas, hemos usado tres barras LED (*BDBL – R(IR)82/16H*), que nos proporcionan un mayor contraste en los filos de corte de las plaquitas. Es importante mencionar que esta máquina no utiliza aceites, lubricantes, o cualquier tipo de sustancia que afecte a la calidad de las imágenes. En la Figura 6 se puede ver un ejemplo de las imágenes obtenidas automáticamente por el sistema de captura.

Después del proceso de adquisición de las imágenes, se llevó a cabo una extracción automática del filo de corte: Primero, se localizó el tornillo central de



Figura 6: Example of the captured images.

las plaquitas usando la transformada circular de Hough (CHT), para detectar círculos con radio entre 40 y 80 píxeles. Este tamaño se fijó teniendo en cuenta el conocimiento de la resolución 2592×1944 . El siguiente paso consistió en aplicar un filtro de Canny para detectar los fillos de corte de las plaquitas. Después de esto, se detectaron las líneas verticales usando la transformada estándar de Hough (SHT). Finalmente, segmentamos la imagen teniendo en cuenta la primera línea vertical encontrada en la parte izquierda de la imagen, que se corresponderá con el filo de corte. El resultado final se muestra en la Figura 7. Al final, hemos obtenido 577 imágenes del filo de corte (301 fillos intactos y 276 con desgaste).

4 Descriptores basados en el contorno.

En este capítulo se proponen dos nuevos métodos basados en la información de contorno. Ambos fueron desarrollados para mejorar la eficiencia y el coste computacional del método desarrollado en (Shu and Wu, 2011b), llamado CPDH. En las siguientes secciones se introduce el método original así como nuestras dos nuevas propuestas.

4.1 CPDH

El método original, Histograma de Distribución de los Puntos de Contorno (CPDH), fue propuesto por Shu et al. (Shu and Wu, 2011b). El primer paso es obtener la representación de los puntos de contorno de la forma en coordenadas polares. Estos puntos son los que se van a usar para describir la imagen y se pueden

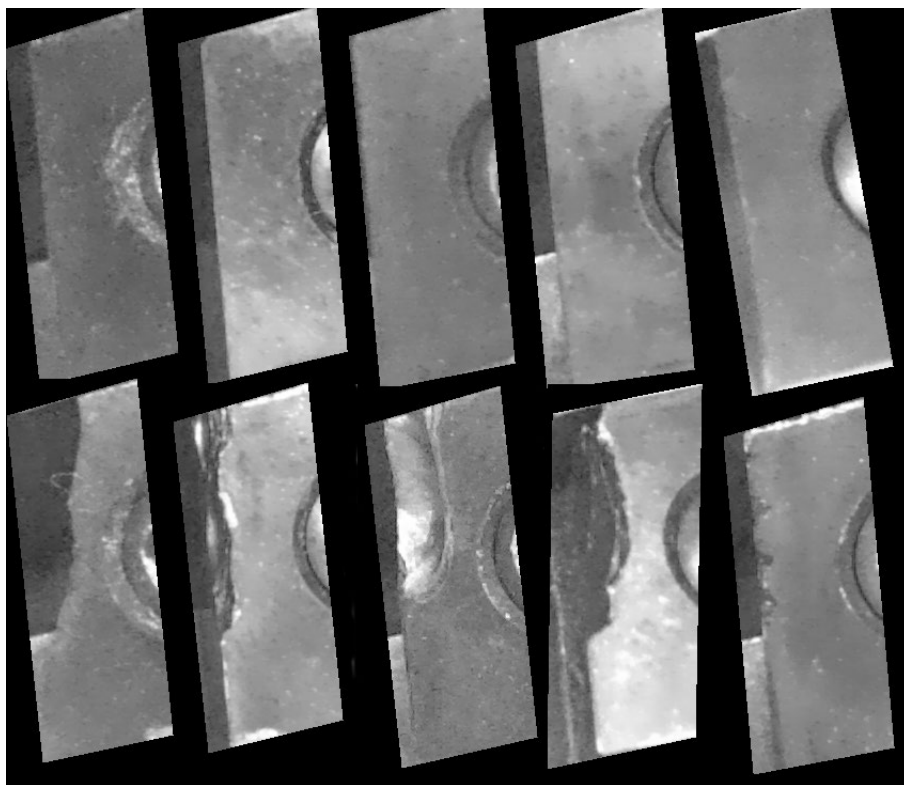


Figura 7: imágenes del filo de corte. En la primera fila se muestran los filos intactos. En la segunda, filos con diferentes niveles de desgaste..

obtener de diferentes maneras. En este caso, los autores emplearon el operador de Canny (Belongie et al., 2002). Una vez que se extraen los puntos de contorno su centroide se fija como origen. Después, se construye una circunferencia circunscrita con ese origen como centro y la región definida por esta circunferencia se divide en porciones mediante circunferencias concéntricas y ángulos igualmente espaciados. El paso final es la construcción del descriptor CPDH teniendo en cuenta el número de puntos que pertenecen a cada porción, y el ángulo y radio de cada división.

4.2 CPDH36R

Para crear el descriptor CPDH36R, nos basamos en el método propuesto por Belongie y su grupo de investigación (Belongie et al., 2002), obteniendo un contorno como el que se muestra en la Figura 8

Este contorno puede ser representado como una colección de puntos del tipo:

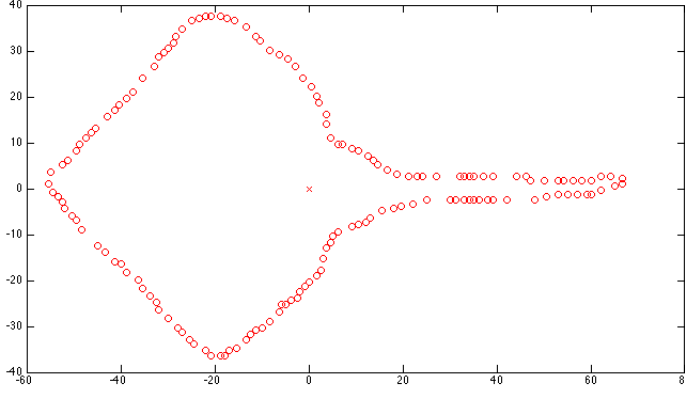


Figura 8: Extracción de los puntos de contorno usando el método de Belongie

$$P = \{(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)\} \quad (1)$$

$$/(x_i, y_i) \in \mathbb{R}^2$$

donde m representa el número de puntos del contorno. Después de esto se obtiene el centroide de los puntos de contorno usando la expresión (2) para obtener la circunferencia circunscrita a la forma.

$$(x_c, y_c) = \frac{\sum_{i=1}^m (x_i, y_i)}{m} \quad (2)$$

El centroide se fija como origen y la colección de puntos, P , se traslada a coordenadas polares mediante la ecuación (3).

$$P = \{(\rho_1, \alpha_1), (\rho_2, \alpha_2), \dots, (\rho_m, \alpha_m)\} \in \mathbb{R}^2 \quad (3)$$

donde

$$\rho_i = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2} \quad (4)$$

y

$$\alpha_i = \begin{cases} \arctan\left(\frac{(y_i - y_c)}{(x_i - x_c)}\right) & \text{if } x > 0, y \geq 0 \\ \arctan\left(\frac{(y_i - y_c)}{(x_i - x_c)}\right) + 2\pi & \text{if } x > 0, y < 0 \\ \arctan\left(\frac{(y_i - y_c)}{(x_i - x_c)}\right) + \pi & \text{if } x < 0 \\ \arctan\left(\frac{\pi}{2}\right) & \text{if } x = 0, y > 0 \\ \arctan\left(\frac{3\pi}{2}\right) & \text{if } x = 0, y < 0 \end{cases} \quad (5)$$

Por tanto, la circunferencia circunscrita tendrá como centro al centroide $C(x_c, y_c)$ y como radio $\rho_{max} = \max(\rho_i), i = 1, 2, \dots, m$.

Una vez hecho esto, el círculo definido por la circunferencia circunscrita se divide en porciones, quedando el área particionada en $u \times v$ porciones donde u es el número de partes en las que se divide ρ_{max} y v indica los sectores en los que se divide el círculo. El resultado de este proceso se puede ver en la Figura 9

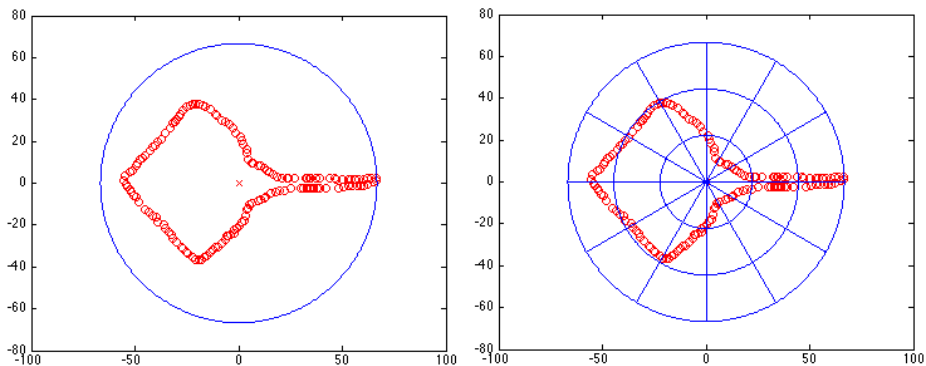


Figura 9: Circunferencia circunscrita al contorno de la imagen (izquierda) y círculo particionado en 36 porciones ($v=12, u=3$) (derecha).

Después de estos pasos previos ya podemos construir el descriptor contando el número de puntos de contorno que corresponden a cada porción. El CPDH original está compuesto por una tripleta por cada porción $H_i = (\rho_i, \theta_i, n_i)$ donde ρ_i indica el radio de la circunferencia concéntrica, θ_i el ángulo y n_i el número de puntos localizados en esa porción r_i . Los autores de CPDH usan el clasificador EMD (Earth Mover Distance) para hacer su descriptor invariante a la rotación. La principal desventaja de este clasificador es su lentitud, por tanto, una de nuestras principales motivaciones a la hora de desarrollar el descriptor CPDH36R fue la necesidad de crear un descriptor que contuviera únicamente las características más relevantes y obtener de esta manera mejores resultados que el método base a la vez que mejoramos el coste computacional.

Nuestra propuesta, CPDH36R, usa el primer y el tercer elemento de las tripletas previamente mencionadas, que son el radio y el número de puntos que están posicionados dentro de cada porción respectivamente, obteniendo el 36SEQ que se muestra en la Figura 11

Posteriormente, los radios son normalizados siguiendo la expresión (6)

$$r_p = \frac{\rho_{1..u}}{\rho_{max}} \quad (6)$$

Donde ρ_{max} es el radio más grande y $p \in [1.,36]$.

CPDH36R se crea multiplicando el número de puntos de cada porción por su radio normalizado correspondiente (el primer elemento de la tripleta), quedándonos así con las características más relevantes para la clasificación, dado que el segundo elemento es constante para todas las imágenes porque el número de divisiones de la circunferencia es siempre el mismo. Ver Figura 10

$n_1 \times r_1$	$n_2 \times r_2$...	$n_{36} \times r_{36}$
------------------	------------------	-----	------------------------

Figura 10: CPDH36R construido multiplicando el número de puntos de cada porción por su correspondiente radio normalizado

Esta propuesta combinada con un método de clasificación como los k vecinos más cercanos (kNN), mejora al método base CPDH en términos de precisión y coste computacional, como se verá en la sección de resultados.

4.3 RCPDH

En este caso también se usa el método desarrollado por Belongie y su grupo de investigación (Belongie et al., 2002) y se siguen los mismos pasos que se han descrito en la sección anterior. Básicamente, RCPDH consiste en hacer el descriptor CPDH intrínsecamente invariable a la rotación para evitar el uso del clasificador EMD. Una vez que se obtiene el descriptor CPDH original, el tercer elemento de cada tripleta, que se corresponde con el número de puntos que hay dentro de la porción, se tiene en cuenta para construir la secuencia de módulo mínimo. Esta pequeña secuencia se llama CPDH36SEQ. Ver Figura 11.

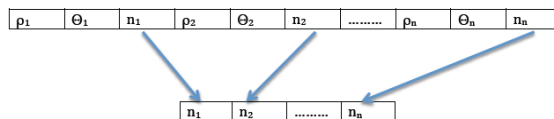


Figura 11: Primer paso para construir RCPDH: extracción de CPDH36SEQ.

Después buscamos la secuencia de módulo mínimo dentro del descriptor CPDH36SEQ desplazando el vector a lo largo de sus 36 posiciones siguiendo la fórmula (7). Ver Figura 12.

$$RCPDH = \text{cat}(CPDH(\rho_i, \theta_i, n_i : \rho_{u \times v}, \theta_{u \times v}, n_{u \times v}), CPDH(\rho_1, \theta_1, n_1 : \rho_{i-1}, \theta_{i-1}, n_{i-1})) \quad (7)$$

donde i satisface la ecuación 8 y $\text{cat}(a, b)$ es el operador concatenador entre a y b .

$$\text{arg}_i \text{ mín}\{\text{cat}(CPDH36(i : u \times v), CPDH36(1 : i - 1)) \forall i \in (1..u \times v)\} \quad (8)$$

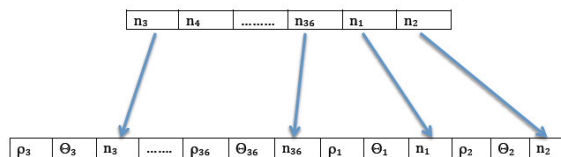


Figura 12: Descriptor RCPDH.

Usando este nuevo descriptor, ya no es necesario el uso de un clasificador que resuelva el problema de la invarianza a la rotación. Como se ve en la sección de resultados, esta propuesta combinada con kNN, también mejora el descriptor CPDH original.

4.4 Reconocimiento de formas

Nuestras propuestas fueron evaluadas usando como clasificadores kNN, kNN por distancia media y kNN ponderado con distancias intersección y cityblock. Las pruebas se realizaron con cuatro conjuntos de imágenes: kimia25, kimia99, MPEG7 y MPEG2. Los resultados obtenidos para cada conjunto de imágenes usando kNN original con $k=1$ se pueden ver en las tablas 1 and 2.

El caso de Kimia25 es el único en el que el mejor resultado se obtiene usando el clasificador EMD debido al pequeño número de muestras, obteniendo un 90 % de acierto, frente a nuestro 76 % con RCPDH y 84 % usando CPDH36R.

En la tabla 1 podemos ver que los resultados son mejores utilizando el conjunto de imágenes kimia99 en lugar de kimia25 y nuestras propuestas CPDH36R y RCPDH en comparación con el método CPDH original. Viendo los resultados en la tabla 1, podemos concluir que nuestra propuesta CPDH36R es la mejor de las tres, por tanto, la utilizamos para probar con los conjuntos de imágenes más desafiantes. En la tabla 2 se puede ver también la eficiencia de nuestro CPDH36R. Usando el conjunto de imágenes MPEG7, una vez más, nuestra propuesta CPDH36R mejora el

método del estado del arte (CPDH) con todas las distancias empleadas. Utilizando el conjunto de imágenes MPEG2, nuestra propuesta también mejora al método base con todas las distancias obteniendo un 91.25 % de acierto en el caso de la distancia cityblock. Aunque en conjuntos de imágenes más simples como kimia25, el método original obtiene mejores resultados que nuestras propuestas, en los conjuntos más complejos, CPDH36R obtiene mejoras significativas con respecto al método original. Es importante señalar que para el conjunto de imágenes MPEG7, el cálculo de la combinación de CPDH con EMD no ha sido posible debido al gran coste computacional y al tamaño del conjunto de imágenes (1400).

Tabla 1: Clasificación Kimia25 y Kimia99 usando CPDH original RCPDH y CPDH36R con $k=1$.

Clasificador	Kimia25			Kimia99		
	CPDH	RCPDH	CPDH36R	CPDH	RCPDH	CPDH36R
kNN Intersección	60 %	64 %	80 %	83.83 %	89.89 %	93.93 %
kNN Cityblock	68 %	76 %	84 %	91.91 %	91.91 %	94.94 %
EMD	90 %	-	-	86.41 %	-	-

Tabla 2: Clasificación MPEG7 y MPEG2 usando CPDH original yCPDH36R con $k=1$.

Clasificador	MPEG7		MPEG2	
	CPDH	CPDH36R	CPDH	CPDH36R
kNN Intersección	50.36 %	82.07 %	67.50 %	91.25 %
kNN Cityblock	59.14 %	82.14 %	74.75 %	91.25 %
EMD	-	-	44.25 %	-

En las Figuras 13 y 14, se pueden ver los resultados con nuestra mejor propuesta y con el método original, usando como clasificador kNN con distancia media y también ponderado para diferentes valores de k y con distancia cityblock. Como se puede ver en todas las figuras, usando kNN original la precisión de recuperación decrece para valores grandes de k . Este problema se soluciona utilizando kNN por distancia media o bien kNN ponderado en los que las variaciones son menores. En la Figura 13, usando el conjunto de imágenes Kimia25, se puede ver que el rendimiento de kNN original empeora para valores altos de k , obteniendo resultados más estables usando el kNN por distancia media o el ponderado. En el resto de conjuntos de imágenes, cuyas tasas de acierto se muestran en las figuras 13 y 14, se ve que los resultados más estables se obtienen finalmente con el método de la distancia media.

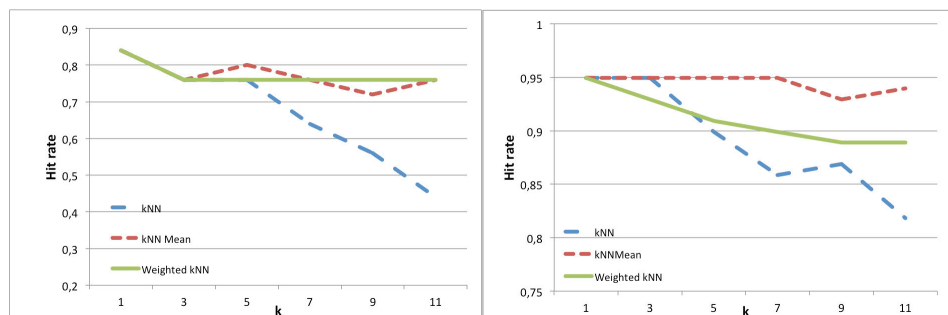


Figura 13: Tasa de acierto para los diferentes métodos kNN aplicados en el descriptor CPDH36R usando los conjuntos Kimia25 (izquierda) y Kimia99 (derecha).

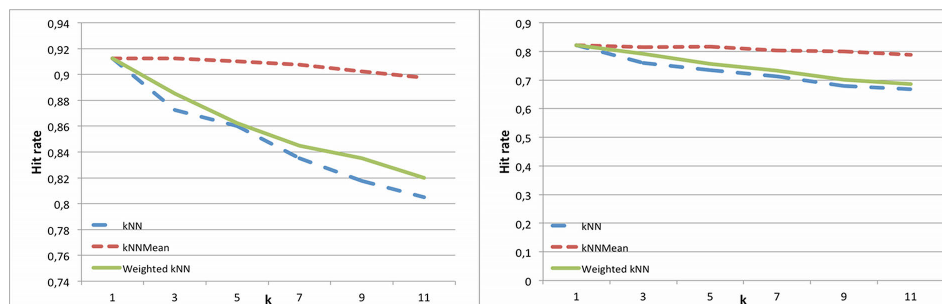


Figura 14: Tasa de acierto para los diferentes métodos kNN aplicados en el descriptor CPDH36R usando los conjuntos MPEG2 (izquierda) y MPEG7 (derecha).

4.5 Conclusiones

En este capítulo se han propuesto dos métodos de recuperación de imágenes: RCPDH y CPDH36R, ambos basados en el descriptor CPDH. El descriptor original usa el clasificador EMD, método lento y pesado. Nuestras propuestas pueden clasificarse usando kNN con métricas clásicas, como distancia euclídea, intersección, cityblock, etc, obteniendo resultados en unos pocos segundos mientras que utilizando el clasificador EMD tardaríamos días. Como nuestra propuesta CPDH36R ofrece mejores resultados que los demás métodos, fue probada con conjuntos de imágenes sencillos y también conjuntos de imágenes desafiantes: kimia25, kimia99, MPEG7 y MPEG2, obteniendo resultados prometedores y alcanzando no solo las más grandes diferencias en cuanto a precisión en el conjunto más complejo, MPEG7, si no también un gran rendimiento en MPEG2 (91.25 %) y kimia99 (94.94 %).

5 Descriptores basados en orientaciones y momentos

En el capítulo 4, propusimos descriptores de contorno para caracterizar el desgaste de la herramienta. Este tipo de descripción tiene limitaciones, ya que el contorno no nos da toda la información que necesitamos acerca del desgaste. Por tanto, con el objetivo de mejorar los resultados obtenidos, se consideró desarrollar descriptores basados en momentos y orientaciones. En este capítulo se proponen dos nuevos métodos basados en momentos y en orientaciones de contorno. Ambos se han llevado a cabo teniendo como método de partida el descriptor ZMEG desarrollado por Anuar y su grupo de investigación (Anuar et al., 2013). La idea de ambas propuestas es combinar descriptores locales con globales. Por una parte, los locales describen regiones pequeñas de la imagen y por otra parte, los globales nos dan información sobre la forma general de la imagen. De esta manera, contamos con las ventajas de ambos. En las secciones siguientes se introduce el método original ZMEG y se presentan nuestras dos propuestas: aZIBO y B-ORCHIZ.

5.1 ZMEG

ZMEG (Anuar et al., 2013), es un descriptor de forma que combina características globales y locales para imágenes binarias. Hace uso de los momentos de Zernike como descriptor global (específicamente el módulo de los 36 primeros coeficientes, hasta orden 10). Es bien sabido que el módulo de los momentos de Zernike de una imagen f tiene la propiedad de ser invariante a la rotación. El descriptor local de forma empleado se llama EGCM (Edge Gradient Co-occurrence Matrix) y se obtiene como sigue: Primero se obtienen los puntos de contorno de la imagen. En este caso, el detector de bordes es el propuesto por Belongie Belongie et al. (2002) que devuelve una muestra de los puntos del contorno uniformemente distribuidos. El resultado de este proceso se puede ver en la Figura 15

El conjunto de puntos del contorno se representa como:

$$P = \{(x_i, y_i) \in \mathbb{R}^2 : 1 \leq i \leq m\},$$

donde m es el número de ellos.

En el segundo paso se calcula el gradiente de cada punto de contorno P mediante la expresión (9).

$$O(x, y) = \arctan \left(\frac{I(x+1, y) - I(x-1, y)}{I(x, y+1) - I(x, y-1)} \right), \quad (9)$$

donde x e y son las coordenadas de la imagen.

Estas orientaciones de gradiente se discretizan en los ocho valores mostrados en la Figura 15(c).

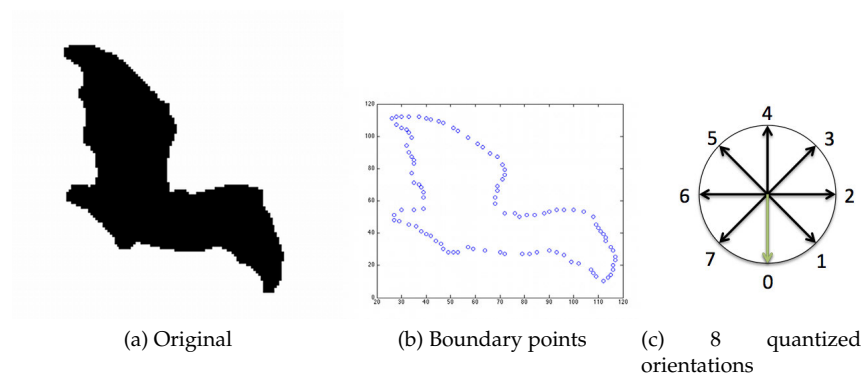


Figura 15: Pasos seguidos hasta obtener las ocho orientaciones cuantizadas.

Como tercer paso del proceso se construye la matriz EGCM teniendo en cuenta solo aquellos puntos, dentro de los ocho vecinos, que formen parte del contorno. La Figura 16 muestra la construcción de la matriz EGCM para un punto.

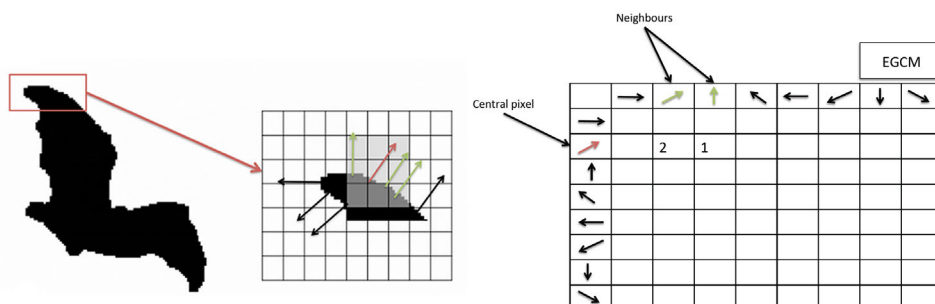


Figura 16: Ejemplo de las orientaciones que se tienen en cuenta para un punto de contorno, mostrado en rojo, para la construcción de EGCM (izquierda). Ejemplo de la construcción de EGCM teniendo en cuenta el punto mostrado en la figura (derecha).

Finalmente, las ocho filas de la matriz EGCM se concatenan dando como resultado un vector de 64 elementos. En el trabajo desarrollado por Anuar (Anuar et al., 2013), los valores del descriptor se normalizan en el rango 0-1.

5.2 aZIBO

El descriptor global usado por aZIBO sigue el mismo método base que en ZMEG pero en este caso se utilizan las imágenes con valores de intensidad entre 0 y 255 (no binarias) y además son redimensionadas a un tamaño de 128×128 para dar mayor

información a los coeficientes de los momentos de Zernike gracias al proceso de interpolación que genera imágenes en escala de grises en vez de imágenes binarias. El descriptor local EGCM se hace invariante a la rotación. Para ello, se busca el par de puntos de contorno b_1 y b_2 cuya distancia euclídea es más grande y después, el ángulo del vector $\vec{b} = b_1\vec{b}_2$ se calcula mediante la ecuación (10).

$$\alpha_{\vec{b}} = \arctan \left(\frac{b_2(2) - b_1(2)}{b_2(1) - b_1(1)} \right), \quad (10)$$

En caso de que haya más de un par de puntos con la misma distancia, se coge el par (b_1, b_2) de manera que el segmento que une los puntos, $\vec{b}_1\vec{b}_2$, sea el más cercano al centroide de la forma. El punto b_1 o b_2 con más puntos de contorno en su vecindario de 21×21 es el que seleccionamos como origen. El tamaño del vecindario se ha determinado experimentalmente. La orientación gradiente del origen, ϕ_d , se coloca en la primera fila y columna de la matriz EGCM y el resto de las orientaciones son desplazadas siguiendo su posición original. Por tanto, si EGCM es $[\phi_1, \phi_2, \dots, \phi_d, \dots, \phi_8]$, entonces IEGCM (que es su versión invariante), será $[\phi_d, \dots, \phi_8, \phi_1, \dots, \phi_{d-1}]$.

Se muestra un ejemplo de todo este proceso en la Figura 17, donde se ve que se obtiene la misma matriz IEGCM para la misma imagen aunque esté rotada.

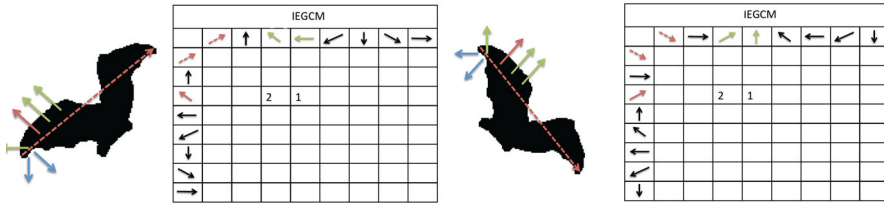


Figura 17: Ejemplo de invarianza a la rotación de IEGCM. La línea roja punteada representa la orientación entre los puntos más alejados del contorno. La línea roja se corresponde con la orientación del gradiente del punto para el que se está calculando el IEGCM. Las flechas verdes son los puntos que se tienen en cuenta (porque están dentro del vecindario) mientras que las azules representan puntos fuera del área de influencia del punto que estamos tratando.

Finalmente se concatenan las ocho filas de IEGCM dando lugar a un vector de 64 elementos. A diferencia del trabajo desarrollado en (Anuar et al., 2013), donde este descriptor es normalizado en el rango 0-1, aZIBO preserva los valores originales para obtener la mayor cantidad de información posible.

5.3 B-ORCHIZ

En esta sección presentamos un descriptor nuevo, que es una extensión de ZMEG y aZIBO, mejorando su efectividad y robustez. B-ORCHIZ usa IBD (Descriptor Invariante del Contorno) como descriptor local, que es una fusión el IEGCM y de la cadena de orientaciones del contorno: BOC. Además, el descriptor global también es invariante a la rotación ya que se usa el módulo de los momentos de Zernike hasta orden diez, de la misma manera que se usaba en el descriptor aZIBO.

La primera componente del descriptor local, IEGCM, fue descrita en la sección anterior. A continuación presentamos la segunda componente, BOC (Boundary Orientatios Chain). BOC es un vector cuyas componentes son las orientaciones de los m puntos de contorno de la forma. Para conseguir la invarianza a la rotación, el primer paso es determinar el primer elemento de esta cadena. El proceso para la extracción de BOC es el siguiente:

1. Como primer paso, se buscan los dos puntos de contorno más alejados b_1 y b_2 , siguiendo el mismo proceso que se usó para encontrar la primera orientación en la construcción de IEGCM.
2. De esos dos puntos, el seleccionado como origen será aquel que tenga más puntos de contorno vecinos en una ventana de 21×21 .
3. Posteriormente se normalizan los valores de orientación para conseguir la invarianza, de manera que la primera orientación, sea cual sea, tiene asignado el valor 0 en la cadena y el resto de orientaciones son desplazadas de acuerdo con esa primera orientación tomando los valores siguientes. Se puede ver un ejemplo en la Figura 18



Figura 18: Ejemplo de construcción de la cadena de orientaciones del borde (BOC) invariante a la rotación. La flecha roja representa la distancia entre los puntos más alejados del contorno. La verde es la orientación del primer punto de la cadena. Los valores se asignan a cada orientación secuencialmente empezando por la flecha verde, la cual colocamos en la parte alta del diagrama como punto de origen (valor 0).

Finalmente, el descriptor invariante es el vector de m elementos que representa la orientación de cada punto de contorno. En esta propuesta, el descriptor local IBD es la concatenación del previamente explicado IEGCM y la cadena BOC. Después de todo este proceso, el descriptor final B-ORCHIZ está formado por la concatenación del descriptor global y el descriptor local IBD.

5.4 Reconocimiento de forma

Se probaron los tres descriptores, B-ORCHIZ, aZIBO y ZMEG con conjuntos de imágenes clásicos. La clasificación con las imágenes de forma fue llevada a cabo usando el clasificador de los vecinos más cercanos (kNN) con distancias coseno, euclídea, intersección, chi cuadrado y cityblock. Hemos seleccionado este tipo de clasificación porque es la más usada en el estado del arte para conjuntos de imágenes de forma (Kimia y MPEG7).

Como se ha dicho previamente, B-ORCHIZ es una combinación de descriptores globales y locales. La diferencia de magnitud entre uno y otro es muy grande, por tanto es necesario llevar a cabo una normalización ponderada para que la clasificación no quede sesgada debido a esas diferencias de magnitud. En nuestro caso, ponderamos ambos descriptores (global y local) usando la ecuación (11).

$$B_{ORCHIZ} = Z \cdot w_g + IBD \cdot w_l, \quad (11)$$

donde Z representa el vector de los 36 módulos de los momentos de Zernike. Los valores de los pesos han sido obtenidos experimentalmente y fijados $w_g = 1$ y $w_l = 5$.

En la Tabla 3 se pueden ver las tasas de acierto obtenidas clasificando con kNN para B-ORCHIZ, aZIBO y ZMEG. A la vista de los resultados, B-ORCHIZ mejora al resto de los métodos en los tres conjuntos de imágenes e independientemente de la métrica de distancia empleada.

Como se puede observar en la primera parte de la tabla 3, la mejor tasa de acierto para el conjunto de imágenes Kimia, 91.92 %, se obtiene describiendo con B-ORCHIZ usando la distancia coseno, mientras que ZMEG y aZIBO obtienen una tasa de acierto de 87,88 %.

Si nos centramos en los resultados del conjunto de imágenes MPEG-2 (ver tabla 5.1, segunda parte), B-ORCHIZ obtiene un 93.75 % de tasa de acierto usando la distancia coseno mientras que aZIBO (con distancia euclídea) y ZMEG (con distancia cityblock) obtienen un 79.50 % y 79.00 % respectivamente. Cabe destacar que en todos los experimentos llevados a cabo se obtienen los mejores resultados con el vecino más cercano ($k = 1$). El rendimiento de la clasificación decrece al tiempo que aumenta el número de vecinos a tener en cuenta en casi todas las pruebas.

Con respecto a los resultados usando el conjunto de imágenes más complejo para

la recuperación, MPEG-7, mostrados en la tercera parte de la Tabla 3, una vez más B-ORCHIZ mejora al resto de los descriptores para cualquier distancia. En este caso, B-ORCHIZ obtiene una tasa de acierto del 89.50%, mientras que aZIBO y ZMEG obtienen un 77.07% y 76.57%, respectivamente.

Tabla 3: Tasas de acierto en % de nuestra propuesta B-ORCHIZ contra aZIBO y ZMEG usando kNN (con diferentes distancias) en los conjuntos de imágenes Kimia-99 (primero), MPEG-2 (segundo) y MPEG-7 (tercero).

	kNN Coseno	kNN Euclidea	kNN Intersección	kNN Chi-cuadrado	kNN Cityblock
B-ORCHIZ	91.92	90.91	90.91	89.90	90.91
aZIBO	85.86	86.87	86.87	86.87	87.88
ZMEG	87.88	85.86	83.84	72.73	83.84
B-ORCHIZ	93.75	93.00	92.25	88.50	92.75
aZIBO	72.25	79.50	75.75	79.25	79.00
ZMEG	75.25	79.00	52.75	61.25	79.00
B-ORCHIZ	89.50	88.29	87.29	84.36	88.21
aZIBO	75.00	75.79	73.29	76.50	77.07
ZMEG	75.93	76.29	38.36	56.79	76.57

Las curvas precision-recall (PR) de los tres descriptores con diferentes valores de k , se muestran en la Figura 19. Podemos ver claramente que el descriptor B-ORCHIZ mejora el rendimiento de los demás en las tres clasificaciones llevadas a cabo.

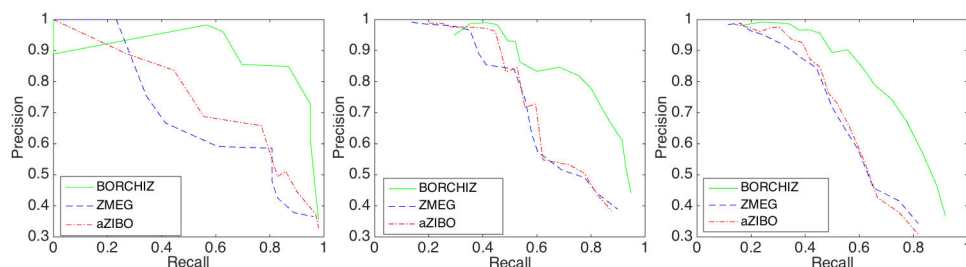


Figura 19: Curvas PR para los conjuntos de imágenes Kimia-99, MPEG-2 y MPEG-7 descritos con B-ORCHIZ, aZIBO y ZMEG.

Estos resultados obtenidos para los conjuntos de imágenes clásicos confirman la efectividad y precisión de B-ORCHIZ e incluso de aZIBO en comparación con el método base propuesto en (Anuar et al., 2013) para propósitos de recuperación de formas.

5.5 Monitorización del desgaste de la herramienta

En el caso del conjunto de imágenes de fillos de plaquitas, utilizamos la técnica de los k vecinos más cercanos con valores de $k=1, 3, 5, 7, 9$ y 11 y seis distancias:

coseno, euclídea, intersección, chi cuadrado, sqdist y cityblock para clasificar la concatenación de los descriptores globales y locales previamente explicados. LLevamos a cabo dos tipos de clasificación. La primera considerando tres clases de desgaste (bajo (L), medio (M) y alto (H)) y la segunda considerando solo dos clases (bajo (L) y alto (H) desgaste).

En el caso del conjunto de imágenes de regiones de plaquitas evaluamos empíricamente la precisión de los tres métodos. Ambos subconjuntos del total de las imágenes (Insert-C e Insert-I) se evaluaron primero de manera independiente y después de forma conjunta para estudiar la posible influencia de los filos de corte verticales en el filo de corte principal. En este caso, y debido al pequeño tamaño de algunas de las regiones, la clasificación fue llevada a cabo utilizando máquinas de vector soporte (SVM) con kernel lineal, polinomial de orden dos e intersección. Se utilizó el algoritmo de los mínimos cuadrados como método de entrenamiento y validación cruzada con diez particiones estratificada en todos los casos.

En las tablas 4 and 5 se muestran las mejores tasas de acierto obtenidas para la clasificación con tres métodos de descripción (aZIBO, Hu y Flusser). Las distancias coseno, chi cuadrado y sqDist presentan siempre los mejores resultados y se puede ver que aZIBO obtiene mejores resultados que los descriptores clásicos, Hu y Flusser, alcanzando tasas de acierto del 60.37% para la clasificación ternaria y 81.13% para la clasificación en dos clases de desgaste.

Tabla 4: Clasificación de las plaquitas usando nuestra propuesta aZIBO y dos descriptores clásicos: Hu y Flusser con valor de k igual a nueve y distancias coseno, chi cuadrado, sqDist y cityblock y etiquetado en tres clases diferentes de desgaste.

	Coseno	ChiCuadrado	SqDist	Cityblock
aZIBO	58.96 %	57.07 %	60.37 %	59.43 %
Hu	55.66 %	56.13 %	59.91 %	58.49 %
Flusser	57.07 %	58.96 %	57.08 %	57.08 %

Tabla 5: IClasificación de las plaquitas usando nuestra propuesta aZIBO y dos descriptores clásicos: Hu y Flusser con valor de k igual a nueve y distancias coseno, chi cuadrado, sqDist y cityblock y etiquetado en dos clases diferentes de desgaste.

	Coseno	ChiCuadrado	SqDist	Cityblock
aZIBO	81.13 %	80.66 %	81.13 %	81.13 %
Hu	76.89 %	79.72 %	79.25 %	80.19 %
Flusser	79.25 %	78.30 %	75.47 %	75.47 %

En la Figura 20 se pueden observar gráficamente y con más detalle los resultados

de aZIBO, Hu y Flusser, para todos los valores de k y seis distancias.

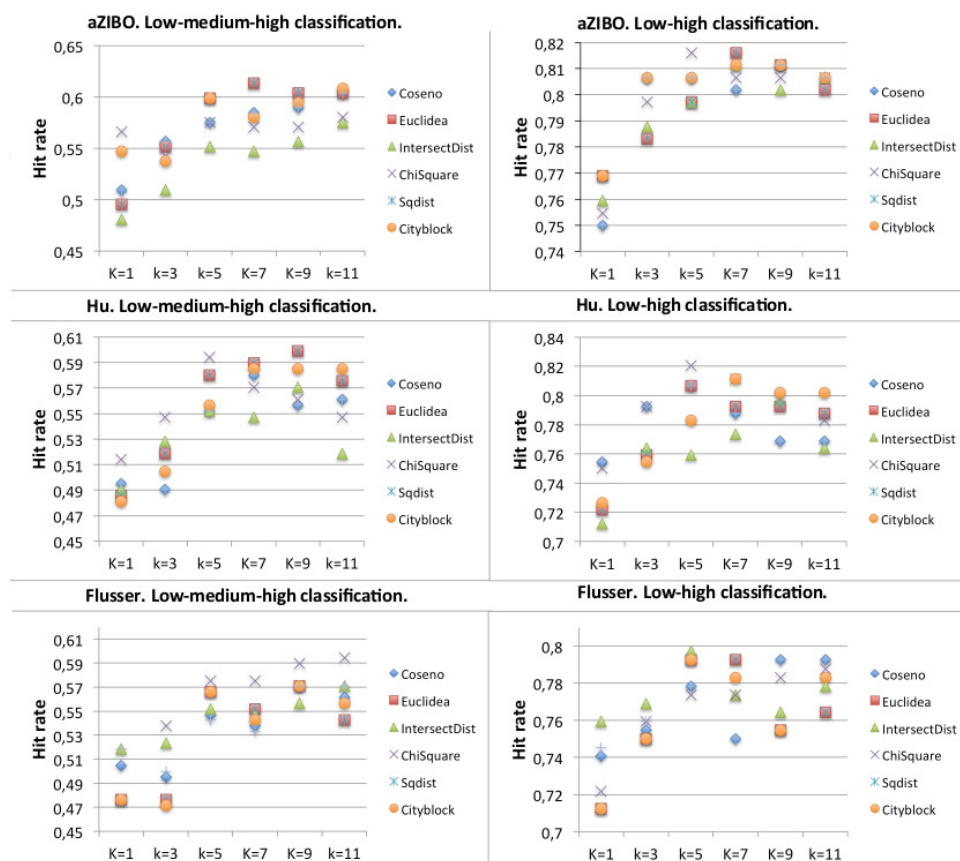


Figura 20: Clasificación con kNN usando seis distancias y diferentes valores de k , para aZIBO, Hu y Flusser y dos categorías de etiquetado del conjunto de imágenes: L-M-H y L-H (3 clases de desgaste y dos clases de desgaste)

Como se puede ver, en la clasificación L-H siempre se obtienen mejores resultados que en la L-M-H aunque proporciona menos información acerca del desgaste. De la misma manera que en el caso anterior, se ve que aZIBO mejora al resto de descriptores en casi todos los casos.

Si nos centramos ahora en el conjunto de imágenes de regiones de plaquitas, para los dos subconjuntos que obtuvimos (Insert-C e Insert-I), podemos ver los resultados obtenidos con nuestras propuestas aZIBO y B-ORCHIZ y el método del estado del arte ZMEG para los dos etiquetados (L-H y L-M-H) en las tablas 6 y 7.

También se llevó a cabo otra clasificación combinando ambos subconjuntos para determinar la influencia de los fillos verticales en los horizontales. Los resultados se

muestran en la tabla 8.

Para el subconjunto Insert-C (Tabla 6), B-ORCHIZ y aZIBO obtienen una precisión del 87.02 % en la clasificación binaria usando el kernel intersección mientras que ZMEG obtiene un 85.58 %. En la clasificación que supone más reto, la clasificación en tres clases, el rendimiento de B-ORCHIZ es más alto que el de los otros dos métodos con un 81.25 %, en contra del 76.92 % y 76.44 % obtenidos con aZIBO y ZMEG, respectivamente. También cabe destacar que en todos los experimentos, el kernel intersección es el que mejores resultados proporciona, mejorando la clasificación L-M-H en un 35.19 % y en un 60.95 % con respecto al kernel lineal y el polinomial respectivamente.

Tabla 6: Resultados de precisión (en %) usando SVM (con kernels lineal, cuadrático e intersección) para las clasificaciones binaria y ternaria y el conjunto de imágenes the Insert-C.

Clasificador	L-H			L-M-H		
	B-ORCHIZ	aZIBO	ZMEG	B-ORCHIZ	aZIBO	ZMEG
SVM-Lineal	70.77	72.12	73.08	60.10	58.17	58.17
SVM-Cuadrático	80.29	75.96	77.88	50.48	49.52	49.04
SVM-Intersección	87.02	87.02	85.58	81.25	76.92	76.44

En el caso del conjunto Insert-I (ver Table 5.5), el comportamiento de todos los métodos es bastante similar al anterior.

En el problema binario B-ORCHIZ mejora al resto de descriptores llegando a obtener una precisión del 88.46 %: un 4.21 % y 5.91 % más alta que las obtenidas con aZIBO y ZMEG respectivamente. En la clasificación L-M-H, ZMEG y aZIBO obtienen tasas de acierto del 79.42 % y 82.14 % respectivamente, mientras que B-ORCHIZ consigue un 82.69 %.

Una vez más, el mejor rendimiento se obtiene utilizando el kernel intersección en todas las pruebas llevadas a cabo.

Tabla 7: Resultados de precisión (en %) usando SVM (con kernels lineal, cuadrático e intersección) para las clasificaciones binaria y ternaria y el conjunto de imágenes the Insert-I.

Clasificador	L-H			L-M-H		
	B-ORCHIZ	aZIBO	ZMEG	B-ORCHIZ	aZIBO	ZMEG
SVM-Lineal	78.02	76.37	76.65	61.54	70.60	72.25
SVM-Cuadrático	76.37	67.58	74.84	61.26	60.16	64.01
SVM-Intersección	88.46	84.89	83.52	82.69	82.14	79.40

El tercer experimento consiste en juntar los dos subconjuntos de imágenes, lo

cual da lugar a una clasificación con más riesgo de fallo debido a las variaciones intra clase que presenta el conjunto de imágenes completo. El objetivo de esta tercera prueba es evaluar la influencia de los fillos verticales en la evaluación final del filo de corte completo.

Los experimentos demuestran (ver tabla 8) que en términos de precisión, la fusión de ambos conjuntos de imágenes no mejora la clasificación, consiguiendo resultados similares a los obtenidos usando el conjunto Insert-C y ligeramente peores que usando Insert-I.

Tabla 8: Resultados de precisión (en%) usando SVM (con kernels lineal, cuadrático e intersección) para las clasificaciones binaria y ternaria y el conjunto de imágenes Insert (conjunto completo).

Clasificador	L-H			L-M-H		
	B-ORCHIZ	aZIBO	ZMEG	B-ORCHIZ	aZIBO	ZMEG
SVM-Lineal	80.42	78.32	79.90	62.59	63.11	66.96
SVM-Cuadrático	81.12	71.33	72.03	72.52	57.69	61.36
SVM-Intersección	87.06	84.44	83.74	80.24	78.85	75.87

En la Figura 21 se pueden ver las curvas PR de la clasificación usando B-ORCHIZ y el resto de descriptores evaluados para los conjuntos de imágenes Insert-C, Insert-I y el conjunto completo Insert. Se puede ver que B-ORCHIZ obtiene los mejores resultados en los tres escenarios evaluados.

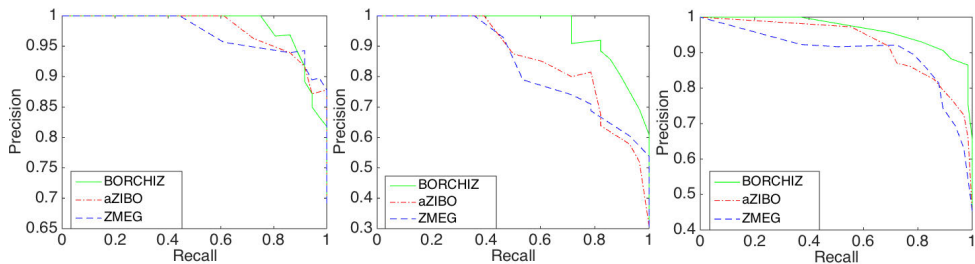


Figura 21: Curvas PR describiendo con B-ORCHIZ, aZIBO y ZMEG los conjuntos de imágenes Insert-C, Insert-I e Insert (de izquierda a derecha).

Finalmente se llevó a cabo un análisis con dendrogramas. Un dendrograma es un grafo estructurado que representa el resultado de un agrupamiento jerárquico entre diferentes clases. Proporcionan información específica sobre la correlación entre grupos, apareciendo en la parte alta o baja del grafo. A medida que nos vamos moviendo hacia la parte superior del dendrograma, mayor es la distancia entre los

diferentes grupos.

Nosotros llevamos a cabo un agrupamiento jerárquico para comparar los resultados de las clases con las generadas por el experto. Utilizamos el método de linkage Ward (Ward, 1963) y el experimento lo realizamos para los tres descriptores evaluados ZMEG, aZIBO and B-ORCHIZ y los tres conjuntos de imágenes. Los dendrogramas generados se pueden ver en la Figura 22.

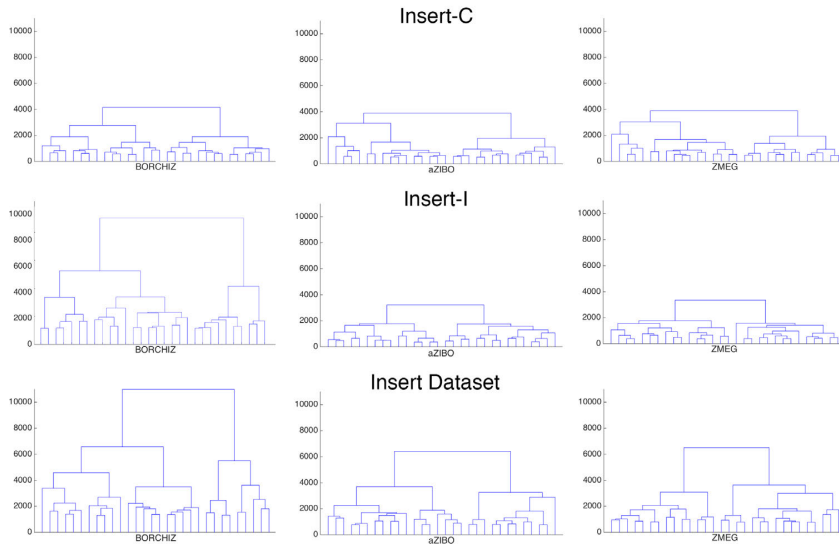


Figura 22: Dendrogramas de agrupamientos obtenidos usando los descriptores B-ORCHIZ (izquierda), aZIBO (centro) y ZMEG(derecha) para el conjunto de imágenes Insert, Insert-C e Insert-I.

Una distancia grande determina grupos separados y por tanto, una clara distinción sobre si una plaquita de corte está dañada o no.

En el conjunto Insert-C la división de las clases es pequeña en cuanto a distancia para todos los descriptores, cerca de 3200 para dos clases y sobre 2000 para tres. Sin embargo, para el conjunto Insert-I subset, la distancia entre dos clases es aproximadamente de 6000 con B-ORCHIZ, pero solo de 2000 con aZIBO y ZMEG. Para tres clases, la distancia entre los agrupamientos más cercanos es de 1500 tanto en aZIBO y ZMEG y de nuevo más alta para B-ORCHIZ (4000). Con esto se concluye que B-ORCHIZ tiene mayor poder discriminativo en el subconjunto de imágenes Insert-I.

En la tercera fila, Figura 22, también muestra que para el conjunto de imágenes completo, B-ORCHIZ vuelve a mostrar mayor poder discriminativo. De acuerdo con el análisis llevado a cabo, se puede concluir que los mejores agrupamientos

para separar plaquitas con diferentes tipos de desgaste, se consiguen con el método propuesto B-ORCHIZ.

También se llevó a cabo una extracción de la forma característica del desgaste para cada clase. El proceso es el siguiente:

1. En primer lugar se alinean todos los fillos al centro de la imagen. Para el conjunto Insert-C, este proceso consiste básicamente en trasladar la región de manera que su fila central aparezca en el centro de la imagen. En el caso del conjunto Insert-I, los fillos izquierdos y derechos tienen diferentes orientaciones dependiendo de la plaquita. Después de trasladarlas al centro de la imagen, tenemos que rotar cada una de estas regiones hasta que queden totalmente verticales.
2. En segundo lugar, todas las imágenes son redimensionadas y se lleva a cabo una suma de sus píxeles, de manera que, al ser imágenes binarias, si la imagen resultante tiene un píxel con valor t , esto quiere decir que hay t imágenes con desgaste en esa posición.
3. Finalmente, la imagen resultante es la media de todas ellas.

Con la obtención de esta imagen representativa para cada clase, se intenta mejorar la comprensión de cómo es el desgaste y también da pistas a los expertos para poder estimar la clase de cada plaquita nueva que tengan que evaluar.

5.6 Conclusiones

En este capítulo se propusieron dos nuevos descriptores de forma basados en ZMEG, aZIBO y B-ORCHIZ. Se llevaron a cabo experimentos con conjuntos de imágenes clásicos como Kimia y MPEG7 para comparar la eficiencia de los tres métodos. Los resultados mostraron que B-ORCHIZ y también aZIBO mejoran los resultados obtenidos con el método original. Una de las principales razones de esta mejora se puede deber a que nuestras propuestas son totalmente invariantes a la rotación mientras que en ZMEG solo son invariantes las características globales pero no las locales.

Para evaluar nuestra propuesta creamos un conjunto de imágenes de 53 herramientas, con cuatro fillos de corte cada una, que fueron etiquetadas por un experto en función del nivel de desgaste, teniendo en cuenta dos y tres clases.

Se utilizó una clasificación supervisada, empleando SVM con los tres descriptores y conjuntos de imágenes y nuestra propuesta B-ORCHIZ obtuvo los mejores resultados en todos los escenarios evaluados. También se llevó a cabo un análisis de agrupamientos que mostró que B-ORCHIZ tiene mayor poder discriminativo a la hora de separar las imágenes en función de la región de desgaste y es el que más información proporciona acerca del estado de dichas regiones.

Estos resultados muestran que sería beneficioso el uso de técnicas de visión para proyectos futuros como métodos para la monitorización del desgaste de la herramienta de corte.

6 Fusión de descriptores de forma y de contorno

En este capítulo se propone un nuevo sistema basado en la combinación de descriptores de forma y de contorno. Para describir el contorno de la forma usaremos el método B-ORCHIZ porque ha demostrado su buen funcionamiento en los experimentos llevados a cabo, como se ha visto en el capítulo anterior, posicionándose como el método que obtiene mejores resultados en la descripción de contorno hasta el momento (García-Ordás et al., 2016). Para la descripción de forma proponemos un descriptor llamado ShapeFeat. Nuestra motivación a la hora de presentar ShapeFeat es que tiene en cuenta las propiedades de la región binaria en vez de solamente la información de contorno a la hora de describir las imágenes. Además, tan solo está formado por diez características, lo cual lo hace muy interesante para los sistemas embebidos o las aplicaciones en las que el almacenamiento sea una restricción. La fusión de B-ORCHIZ con este método se llevó a cabo usando tres técnicas diferentes de fusión y se consiguió mejorar el rendimiento que obtenía cada descriptor por separado. Debido a que cada uno de los descriptores tiene en cuenta diferentes características de las imágenes, la combinación de ambos hace que el poder discriminativo entre formas sea mayor.

6.1 ShapeFeat

Nuestra propuesta ShapeFeat tiene en cuenta diez características extraídas de la región binaria de cada imagen. Nos proporciona información útil sobre la imagen que generalmente no se puede obtener con descriptores de momentos como Hu, Flusser, Zernike, que están más centrados en la información de contorno dejando de lado otras características importantes como el área de la región, la homogeneidad,...

Para la descripción empleamos 10 características extraídas de la región binaria: "ConvexArea", "Eccentricity", "Perimeter", "EquipDiameter", "Extent", "FilledArea", "FilledImage", "MinorAxisLength", "MajorAxisLengthz", "Solidity".

- Área convexa (ConvexArea): número de píxeles del polígono convexo más pequeño que contiene a la región. Una vez que tenemos las coordenadas del polígono más pequeño $((x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n))$, se colocan en forma de determinante. Las coordenadas se cogen en sentido de las agujas del

reloj alrededor de la región empezando y terminando en el mismo punto.

$$Area = \frac{1}{2} \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \dots & \dots \\ x_n & y_n \\ x_1 & y_1 \end{vmatrix} = \frac{1}{2} [(x_1y_2 + x_2y_3 + x_3y_4 + \dots + x_ny_1) - (y_1x_2 + y_2x_3 + y_3x_4 + \dots + y_nx_1)] \quad (12)$$

- **Excentricidad (Eccentricity):** escalar que especifica la excentricidad de la elipse que tiene los mismos segundos momentos centrales que la región (Figura 23). La excentricidad es el ratio de la distancia entre el foco de la elipse y la longitud del eje mayor y su valor está comprendido entre 0 y 1 (0 y 1 son los casos degenerados: una elipse con excentricidad 0 es un círculo y con excentricidad 1 es un segmento).

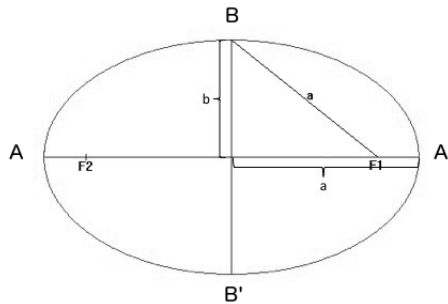


Figura 23: Elipse con centro O . Las coordenadas del foco son $F'(-c, 0)$ y $F(0, c)$

Para cada punto de la elipse: $\bar{P}F + \bar{P}F' = 2a$.

Esta expresión resulta en:

$$\sqrt{(x - c)^2 + y^2} + \sqrt{(x + c)^2 + y^2} = 2a \quad (13)$$

De donde obtenemos:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (14)$$

Entonces, la excentricidad se calcula:

$$\sqrt{1 - \frac{b^2}{a^2}} \quad (15)$$

- **Perímetro (Perimeter):** distancia alrededor de la forma bidimensional. Se corresponde con el número de puntos del contorno. Se representa como $\sum_{i=1}^n P_i$ donde n es el número de puntos de contorno y P_i representa al punto de contorno i .
- **Diámetro equivalente (Equivalent diameter):** escalar que especifica el diámetro del círculo que tenga el mismo área que la región. Se computa como $\sqrt{\frac{4A}{\pi}}$, donde A es el área.
- **Extensión (Extent):** escalar que especifica el ratio de píxeles en la región, y píxeles en la imagen completa.
- **Área rellenada (FilledArea):** número de píxeles de la imagen después de haber rellenado huecos pequeños.
- **Longitud del eje menor (MinorAxisLength):** longitud del segmento $\overline{BB'}$ que se ve en la Figura 23 (2b).
- **Longitud del eje mayor (MajorAxisLength):** longitud del segmento $\overline{AA'}$ que se ve en la Figura 23 (2a).
- $R = \frac{2b}{2a}$, donde $2b$ y $2a$ son las longitudes del eje menor y mayor, respectivamente, de la elipse que tiene los mismos segundos momentos centrales que la región.
- **Solidez (Solidity):** escalar que indica la proporción de píxeles del menor polígono convexo que puede contener la región y que están al mismo tiempo en la región. Se computa como $\frac{Area}{ConvexArea}$.

En la Figura 24 se puede ver información relevante que se usa en el método de descripción. En (a), se ven el semieje menor y mayor de la elipse que tienen los mismos segundos momentos centrales que la región. Esta información también se tiene en cuenta para calcular la excentricidad y el parámetro R . En (b) podemos ver el recorte (crop) de toda la ventana de la imagen que se usa para obtener la propiedad de extensión simplemente calculando el ratio entre el número de píxeles en la imagen y el número de píxeles blancos (píxeles de la región). En (c), se puede ver el polígono convexo más pequeño. Es necesario para extraer el área convexa y la solidez, que es el ratio entre el número de píxeles blancos en (b) y el número de píxeles blancos en (c).

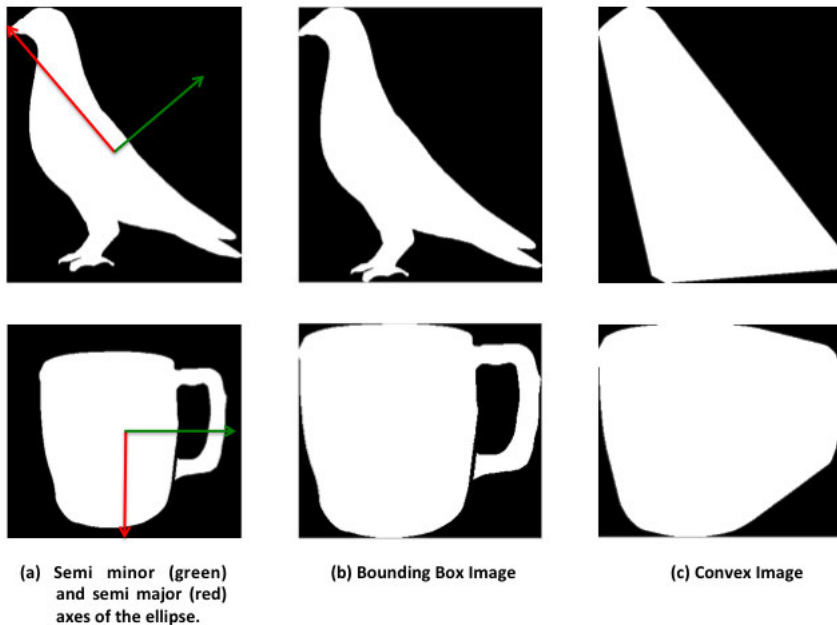


Figura 24: Dos ejemplos de extracción de características de ShapeFeat. Con la información mostrada en la primera columna podemos extraer la excentricidad y el parámetro R y con la mostrada en la segunda, podemos extraer la extensión (extent). Finalmente, el polígono convexo más pequeño mostrado en la tercera columna, es necesario para extraer el área convexa (ConvexArea) y la solidez (Solidity).

6.2 Combinación de descriptores de forma y contorno

En este trabajo exploramos la combinación de características de contorno y de forma para mejorar el sistema de monitorización del desgaste. Podemos destacar tres conocidas estrategias para combinar los descriptores: fusión temprana, fusión intermedia y fusión tardía. A continuación se describe con más detalle cada una de ellas (Ver figura 25).

- Fusión temprana: combina descriptores de diferentes técnicas y produce un solo vector de características para el clasificador. El proceso incrementa el tamaño del vector de características pero es uno de los métodos más rápidos y sencillos de fusión. En nuestro caso, hemos usado la concatenación simple de vectores, por tanto, la entrada al clasificador está compuesta por la concatenación de las características de B-ORCHIZ y ShapeFeat.

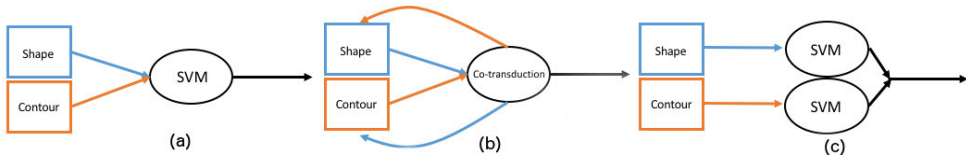


Figura 25: (a) Fusión temprana para descriptores de forma y contorno. (b) Fusión intermedia usando el método de Co-transduction. (c) Fusión tardía usando la media de Bayes.

- **Fusión intermedia:** Cotransduction es un sistema de recuperación desarrollado por Xiang bai et al (Bai et al., 2012). Su objetivo fue desarrollar un algoritmo capaz de fusionar diferentes medidas de similitud para hacer una recuperación de formas más robusta a través de un marco de aprendizaje semi supervisado. El método está inspirado en el algoritmo de co-training (Blum and Mitchell, 1998). En nuestro caso el método se ha adaptado para llevar a cabo una clasificación en vez de una recuperación. Hemos utilizado el método de los k vecinos más cercanos. Además, hemos combinado dos descriptores diferentes en vez de dos medidas de similitud para poder tener en cuenta tanto las características obtenidas con ShapeFeat como con B-ORCHIZ. Dada una imagen de una forma como entrada de prueba, el algoritmo recupera las formas más parecidas a la de entrada usando una medida y luego asignándolas a otro conjunto para la otra medida. Tanto el algoritmo de Co-training como el método de Co-Transduction, están ampliamente explicados en (Bai et al., 2012).
- **Fusión tardía:** utiliza varios clasificadores para determinar la salida en vez de solo uno como hace la fusión temprana. Estos métodos tratan de combinar las puntuaciones de las predicciones, que indican la probabilidad de que una muestra pertenezca a la clase positiva. Aunque es muy sencillo, está comprobado que este método mejora la eficiencia de cada descriptor por separado. En este trabajo hemos usado la media simple de Bayes (Ruta and Gabrys, 2000; Bostrom, 2007) como método para obtener la distribución de probabilidades de los clasificadores fusionados entrenados usando descriptores de contorno (B-ORCHIZ) y de forma (ShapeFeat):

$$P_{LATEFUSION}(x \in C|x) = \frac{P_{SHAPEFEAT}(x \in C|x) + P_{BORCHIZ}(x \in C|x)}{2}, \quad (16)$$

donde x es una imagen y C es una de las posibles clases a las que puede pertenecer. $P_{SHAPEFEAT}$ y $P_{BORCHIZ}$ son las distribuciones de probabilidad para las diferentes clasificaciones llevadas a cabo.

Este método tiene en cuenta la información proporcionada por ambas

clasificaciones y determina la decisión basándose en la media de todas ellas, lo cual nos permite evitar problemas derivados de clasificaciones poco fieles.

6.3 Monitorización del desgaste de herramientas

Para poder determinar el desgaste del filo completo y también la influencia de los filos incompletos, dividimos el conjunto de imágenes en dos: (a) Insert-C con los filos completos que se corresponde con el desgaste horizontal y (b) Insert-I, que se corresponde con los filos incompletos, representados con orientación vertical en las imágenes.

La clasificación para estos tres conjuntos de imágenes se realiza con el clasificador SVM con kernel intersección. Hemos usado validación cruzada Monte-Carlo con divisiones aleatorias del conjunto de datos en entrenamiento y prueba. Para cada división el modelo se ajusta con los datos de entrenamiento y se lleva a cabo una predicción con los datos de prueba. La estimación del rendimiento corresponde a la media de las iteraciones que se hayan hecho (20 en nuestro caso). La ventaja de este método es que la proporción de datos para entrenamiento y para prueba no depende del número de iteraciones. En nuestro trabajo hemos usado una proporción de 75% - 25% para entrenamiento - prueba respectivamente. Los resultados están calculados en términos de precisión. Hemos combinado los dos mejores descriptores (B-ORCHIZ y ShapeFeat) usando fusión temprana, intermedia y tardía y los resultados se han comparado con los obtenidos con cada uno de los descriptores por separado y también con otros descriptores clásicos. El clasificador Co-transduction utiliza k-NN con valores de k: 3, 7, 9 y 11 obteniendo los mejores resultados para k=3 en todos los casos.

6.3.1 Resultados sin fusión para el conjunto de imágenes de regiones de plaquitas

En primer lugar, realizamos una evaluación del método presentado en este trabajo (ShapeFeat) y todas las otras propuestas basadas en contorno: B-ORCHIZ, ZMEG y aZIBO.

En la tabla 9, mostramos los resultados en términos de precisión. Como se puede ver, B-ORCHIZ mejora en todos los conjuntos de imágenes y en todas las clasificaciones a los descriptores ZMEG y aZIBO. La mejora para el conjunto de imágenes completo es más del 3% en la clasificación binaria y casi el 2% en la ternaria con relación al mejor descriptor de la literatura (aZIBO). Por otra parte, estos resultados son mejorados por ShapeFeat en casi todos los experimentos. Como se puede ver, la mayor mejora se obtiene con el conjunto de imágenes Insert-C con un incremento de más del 7.29% de precisión. En los demás casos, el comportamiento es bastante similar en comparación con los métodos de contorno.

Tabla 9: Resultados de la clasificación en % de ZMEG, aZIBO, B-ORCHIZ y ShapeFeat usando SVM con kernel intersección para los conjuntos de imágenes Insert-C, Insert-I (de izquierda a derecha) y para dos y tres niveles de desgaste.

	Complete		Insert-C		Insert-I	
	L-H	L-M-H	L-H	L-M-H	L-H	L-M-H
ZMEG	83.74	75.87	85.58	76.44	83.52	79.40
aZIBO	84.44	78.85	87.02	76.92	84.89	82.14
B-ORCHIZ	87.06	80.24	87.02	81.25	88.46	82.69
ShapeFeat	88.70	80.67	93.37	81.35	88.41	84.12

6.3.2 Resultados para características de contorno y forma fusionadas para el conjunto de imágenes de regiones de plaquitas

B-ORCHIZ (contorno) y nuestro ShapeFeat (forma) son, por el momento, los métodos con los que se consiguen mejores resultados en el escenario de la monitorización del desgaste de la herramienta. La gran diferencia que existe entre los dos métodos hace muy interesante la idea de fusionarlos. Hemos explorado tres técnicas de fusión: temprana, intermedia y tardía.

La fusión temprana se lleva a cabo concatenando B-ORCHIZ y ShapeFeat para crear un nuevo vector de características que será usado como entrada para el clasificador SVM. La fusión intermedia se lleva a cabo utilizando el clasificador previamente explicado Co-Transduction. Por último, la fusión tardía se implementa combinando las puntuaciones del clasificador SVM para B-ORCHIZ y ShapeFeat con el fin de determinar una respuesta final.

En la figura 26 se muestran los resultados para cada método de fusión. Hemos incluido también los resultados por separado de B-ORCHIZ y ShapeFeat con el fin de poder compararlos con facilidad. Como se puede ver, el método de fusión temprana obtiene unos resultados similares a los que obtienen los métodos por separado. Esto puede deberse a la gran descompensación de tamaño entre ambos descriptores, haciendo que B-ORCHIZ predomine sobre ShapeFeat. Co-Transduction y fusión tardía son invariantes al número de características que tenga cada descriptor porque la fusión se lleva a cabo después del paso de clasificación. En casi todos los experimentos, la fusión tardía muestra mejores resultados que el resto de las técnicas. Sin embargo, en la clasificación binaria, en el caso del conjunto Insert-C, la gran diferencia de precisión que existe entre el descriptor de forma y de contorno, condiciona mucho los resultados de la fusión temprana y tardía. Mientras que la fusión temprana y tardía se ven influenciadas por las precisiones de los descriptores, en el caso de Co-Transduction cada descriptor es mejorado por el otro. Sin embargo, a pesar de estas diferencias de tamaño entre los

descriptores, se obtienen los mejores resultados empleando fusión tardía.

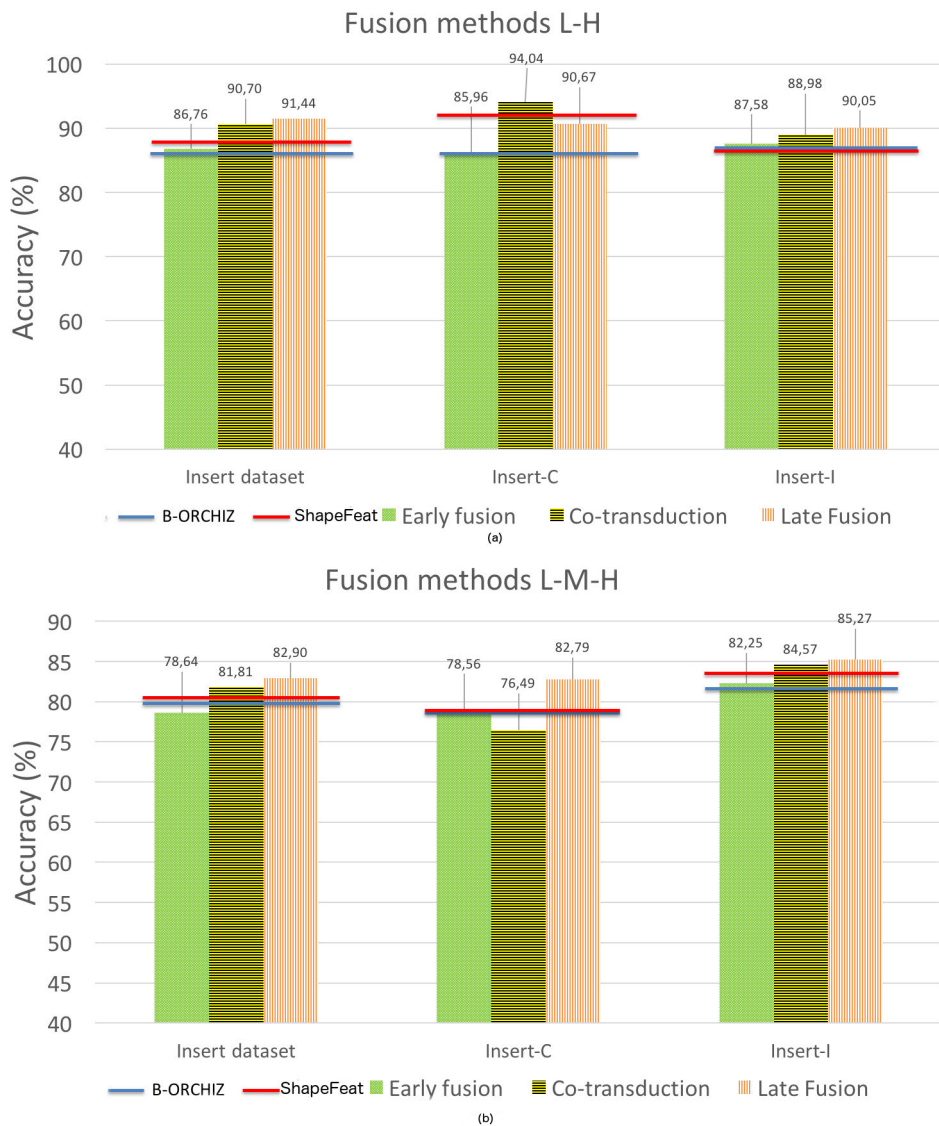


Figura 26: Resultados para la clasificación binaria (a) y ternaria (b) usando la metodología de fusión con B-ORCHIZ y ShapeFeat. Además, se muestra también los resultados alcanzados por cada descriptor por separado, representados como líneas rectas en la gráfica.

6.3.3 Descriptores de contorno+región vs descriptores clásicos

También describimos nuestro conjunto de datos con otros descriptores como BCF (Bag of Contour Fragments), HOG (Histogram of Oriented Gradients) y SC (Shape Context) (Wang, Feng, Bai, Liu and Latecki, 2014; Dalal and Triggs, 2005; Belongie et al., 2002), y en la tabla 10, se puede ver la superioridad de nuestro método combinado comparándolo con los métodos del estado del arte obteniendo mejoras de más del 82 % respecto a SC, más del 35 % respecto a HOG y 31 % para el método BCF.

Tabla 10: Resultados de precisión en % de la combinación de ShapeFeat con B-ORCHIZ y otros descriptores como BCF (Bag of contour fragments), Hog (Histogram of oriented gradients) y SC (Shape Context)

	Complete		Insert-C		Insert-I	
	L-H	L-M-H	L-H	L-M-H	L-H	L-M-H
Late fusion	91.44	82.90	90.67	82.79	90.05	85.27
BCF	76.76	63.11	77.88	49.52	80.77	74.45
HOG	76.80	60.35	80.19	51.06	76.65	66.15
SC	54.58	45.45	68.75	41.83	69.51	48.08

6.4 Conclusiones

En este capítulo se ha presentado un nuevo método para la evaluación automática del nivel de desgaste de plaquitas en procesos de fresado. Evaluando los estudios más recientes relacionados con este campo y con la extracción de características, B-ORCHIZ ha resultado ser la mejor técnica hasta el momento. Este método está basado en la descripción del contorno de la imagen. Con el fin de evaluar otras posibilidades, propusimos el uso de diez características de forma extraídas de la región binaria de la imagen.

Teniendo en cuenta los buenos resultados que se obtuvieron con los descriptores de forma y de contorno, propusimos combinar ambos métodos. La fusión se llevó a cabo usando tres métodos diferentes: fusión temprana, fusión intermedia y fusión tardía. Los resultados obtenidos fueron un 91.44 % de precisión para el dataset completo en la clasificación binaria y un 82.90 % en la ternaria utilizando la técnica de fusión tardía.

7 Descriptor basado en parches

7.1 Metodología

Como ya mencionamos a la hora de proponer los métodos anteriores, los costes derivados del reemplazamiento de las herramientas en el momento inadecuado son muy significativos. Por este mismo motivo, hemos propuesto un sistema portable basado en análisis de textura. En primer lugar, el sistema toma imágenes en escala de grises de las plaquitas del cabezal de corte. Después, se detectan y extraen los fillos de corte automáticamente. A continuación, cada región del filo de corte se divide en parches o regiones y se clasifica cada una de ellas con descriptores de textura, en este caso, basados en LBP. Finalmente, cada uno de los parches descritos se clasifica utilizando Máquinas de Vector Soporte (SVM) y en función del porcentaje de parches de una plaquita clasificados como desgastados, se toma la decisión final acerca del desgaste de esa plaquita.

7.1.1 Configuración de las regiones

Para este problema de descripción de texturas, utilizamos el conjunto de imágenes de alta resolución explicado previamente. Se aborda el problema dividiendo la región del filo de corte en diferentes parches para evaluar el nivel de desgaste de cada uno. Al final, el número de parches desgastados determinarán el nivel de desgaste de la plaquita. Hemos llevado a cabo diferentes evaluaciones en función de la división en parches que se utilice. (Ver Figura 27).

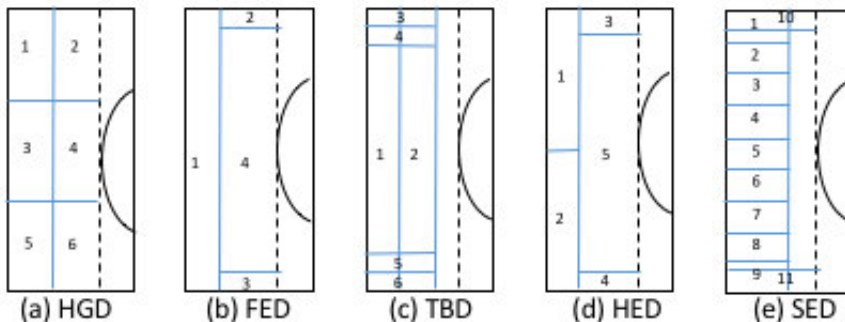


Figura 27: Todas las diferentes configuraciones de división del filo

- División homogénea (HGD): La primera alternativa es la única que tiene un tamaño uniforme de los parches. En este caso se divide el filo de corte de la imagen en 3×2 parches. Con este método intentamos crear parches del mismo tamaño con el objetivo de dar la misma importancia a todas las regiones de

las plaquitas en la etapa de clasificación. Se puede ver con más detalle en la Figura 27 (a).

- División del filo completo (FED): La segunda alternativa está basada en la idea básica de evaluar cada filo de corte independientemente. En la Figura 27 (b) podemos ver el esquema de la división. Para cada imagen se extraen cuatro parches diferentes: el filo de corte principal, dos parches horizontales que serían la parte alta y baja de la plaquita, que normalmente suele tener algún tipo de erosión, y finalmente la región interior de la plaquita. En esta propuesta todos los parches son independientes e incluyen toda la información posible.
- División en dos bandas (TBD): En la tercera alternativa, que se puede ver en la Figura 27 (c), hemos elegido seis parches por cada imagen. El primero, como hicimos en FED, contiene la información del filo principal de corte. Sin embargo, en este caso, también hemos seleccionado un parche del mismo tamaño que el primero que está colocado cuando termina éste, tratando de descubrir regiones de alto desgaste localizadas no solo en el primer parche si no también después del borde de esta primera división. Luego hemos hecho lo mismo con los parches horizontales. Hemos extraído dos parches diferentes en la parte de arriba y otros dos en la parte de abajo. En este caso, los horizontales y verticales tiene partes superpuestas.
- División del filo por la mitad (HED): La cuarta propuesta, representada en la Figura 27 (d), tiene la misma configuración de parches que la primera (FED) pero dividiendo el parche principal, que es el que contiene la información del filo, en dos partes. La primera contiene información del filo principal de corte localizado en la parte de arriba y la segunda contiene la información de la parte inferior. Con esta división intentamos encontrar regiones pequeñas de desgaste que pueden quedar mal clasificadas en la primera propuesta debido a que estén localizadas en solo una parte del parche vertical.
- Pequeñas divisiones del filo (SED): Nuestra última propuesta trata de mezclar algunas ideas usadas en los otros métodos. Como hicimos en TBD, queremos extraer información de la región de desgaste pero en vez de usar dos barras verticales, decidimos usar una región más ancha. Además, tenemos también en cuenta la idea del método HED que divide el parche vertical en dos regiones diferentes pero en este caso hemos querido ir un poco más lejos dividiendo esta región vertical en nueve subregiones. Su representación se puede ver en la Figura 27 (e). También hemos evaluado la parte alta y baja como hicimos en el método FED pero empezando desde la esquina del filo principal, lo que provoca un solapamiento entre estos parches y las dos subregiones de los extremos del parche vertical.

7.1.2 Descriptores de textura

Cada región del filo de corte obtenida, se divide en parches o regiones y posteriormente se clasifica cada una de ellas con descriptores de textura. En este caso, hemos utilizado LBP, ALBP, CLBP y LBPV, que se describen a continuación.

- LBP: Local Binary Pattern (LBP) Ojala and Pietikainen (1996) es un operador de texturas para imágenes en escala de grises. En la ecuación 17 se expresa el concepto de LBP matemáticamente: todos los vecinos que tienen valores más altos o iguales que el valor del píxel central, se multiplican por uno y el resto por cero. Después de esto, se extrae el LBP para cada píxel, sumando todos estos valores. Ver Figura 28 como ejemplo de extracción de un patrón para un píxel.

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p, \quad s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \quad (17)$$

donde g_c es el valor del píxel central, g_p el valor de su vecino, p , P el número de vecinos y R el radio del vecindario.

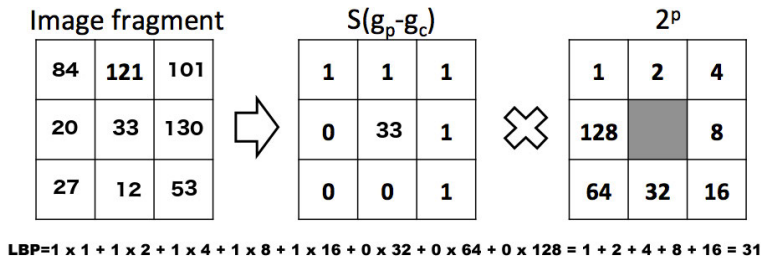


Figura 28: Proceso de extracción de LBP para un píxel en escala de grises usando un vecindario de radio 1 y 8 orientaciones diferentes.

Aunque LBP es invariante a la escala, el hecho de rotar un patrón binario nos daría un descriptor LBP diferente. Por este motivo Ojala y su grupo definieron una nueva ecuación 18 para conseguir la invarianza a la rotación asignando un identificador único a cada LBP rotado.

$$LBP_{P,R}^{ri} = \min \{ROR(LBP_{P,R}, i) | i = 0, 1, \dots, P - 1\} \quad (18)$$

donde $ROR(x, i)$ lleva a cabo un desplazamiento de los bits circular a la derecha en el bit número P y x i veces. De esta manera, por ejemplo los patrones 11001011 y 00101111 nos devolverán el mismo valor de salida.

- ALBP: ALBP es otro descriptor basado en LBP propuesto por Guo y su grupo de investigación Guo, Zhang, Zhang and Zhang (2010). Es adaptativo y está motivado por la falta de información de las orientaciones en el LBP original. Para hacer más robusta la correspondencia contra cambios de estructuras espaciales locales, se tiene en cuenta la media orientada y también la desviación estándar de las diferencias absolutas. Para minimizar las variaciones de la media y la desviación estándar, Guo propone un esquema que minimiza la diferencia direccional $|g_c - w_p * g_p|$ a lo largo de las diferentes orientaciones añadiendo el parámetro w . La función objetivo se define como sigue:

$$w_p = \arg w \min \left\{ \sum_{i=1}^N \sum_{j=1}^M |g_c(i, j) - w \cdot g_p(i, j)|^2 \right\}, \quad (19)$$

donde w_p es el peso usado para minimizar las diferencias direccionales y N y M son el número de filas y columnas de la imagen respectivamente. Cada peso w_p se estima a lo largo de una orientación $2p\pi/P$ para la imagen completa.

En su trabajo, Guo et al. usaron la técnica de los mínimos cuadrados para obtener el vector de pesos óptimos (w). Después de todo esto, el método ALBP se define de la siguiente manera:

$$ALBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - w_p \cdot g_c) 2^p, \quad s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}. \quad (20)$$

- CLBP también fue propuesto por Guo y su grupo de investigación Guo, Zhang and Zhang (2010a) para generalizar y completar al clásico LBP. En este caso la región se representa por su píxel central y un signo de diferencia local (transformada de magnitud llamada LDSMT). LDSMT descompone la imagen en dos componentes: la diferencia de signos y la diferencia de magnitudes. La diferencia de signos (CLBP_S) se calcula de la misma manera que el histograma de LBP y la diferencia de magnitudes (CLBP_M) se define en la ecuación 21. Como ambas tienen carácter binario, pueden ser combinadas para finalmente formar el descriptor CLBP.

$$CLBP_M_{P,R} = \sum_{p=0}^{P-1} t(m_p, c) 2^p, \quad t(x, c) = \begin{cases} 1 & \text{if } x \geq c \\ 0 & \text{if } x < c \end{cases} \quad (21)$$

donde c es un umbral determinado de manera adaptativa.

Finalmente, CLBP se obtiene concatenando o mezclando los dos operadores.

- LBPV: LBPV Guo, Zhang and Zhang (2010b), es otra propuesta de Guo et al

que consiste en combinar LBP con un método de distribución de contrastes. Primero se calcula el LBP uniforme para toda la imagen. Después se usa la varianza de la imagen como peso adaptativo para ajustar la contribución del código LBP en el cálculo del histograma. El histograma LBPV se computa de la siguiente manera:

$$LBPV_{P,R}(k) = \sum_{i=1}^N \sum_{j=1}^M w(LBP_{P,R}(i, j), k), k \in [0, K] \quad (22)$$

donde k representa cada bin del histograma, K es el valor máximo de LBP y w se define como:

$$w(LBP_{P,R}(i, j), k) = \begin{cases} VAR_{P,R}(i, j), & LBP_{P,R}(i, j) = k \\ 0 & \text{otherwise} \end{cases} \quad (23)$$

donde $VAR_{P,R}$ es la varianza del vecindario.

$$VAR_{P,R} = \frac{1}{P} \sum_{p=0}^{P-1} (g_p - u)^2 \quad (24)$$

donde u es la media de los vecinos: $u = 1/P \sum_{p=0}^{P-1} g_p$.

7.2 Monitorización del desgaste de la herramienta

7.2.1 Configuración de los experimentos

Nuestro conjunto de imágenes a alta resolución está compuesto por 577 imágenes del filo de corte divididas en dos subconjuntos: entrenamiento y prueba. El conjunto de entrenamiento contiene el 70 % de las imágenes mientras que el de prueba contiene el 30 %.

Hemos dividido cada imagen del conjunto de entrenamiento en parches que posteriormente han sido etiquetados en función del desgaste. El objetivo de este experimento es generar un modelo para realizar la clasificación individual de los parches de una imagen para, basándose en sus predicciones, decidir el nivel de desgaste de la herramienta. Hacer estos recortes manuales evita una incorrecta extracción de parches que podría deteriorar el correcto funcionamiento del clasificador.

Después de hacer los recortes obtuvimos 896 parches donde 466 se etiquetaron como intactos y 430 como dañados. Entrenamos un modelo SVM con estas imágenes y seleccionamos como kernel el intersección.

En la fase de prueba obtuvimos los parches para cada imagen usando los diferentes métodos de división 27. Todos los parches pertenecientes a una imagen se describieron usando LBP y después fueron clasificados con SVM para determinar si las etiquetas pertenecían a la clase de desgastadas o de intactas. Una vez que tuvimos todos los parches clasificados, calculamos la proporción de parches desgastados en la imagen.

Para determinar si una imagen es intacta o dañada establecemos un umbral que clasifica las imágenes que contienen un número alto de parches etiquetados como desgastados en la clase de las desgastadas. El umbral puede variar entre uno y el número de regiones en la que está dividida la imagen dependiendo del método de división elegido. Cuanto más alto sea el umbral más estricto será el clasificador a la hora de clasificar una plaquita como dañada. En todos los experimentos que llevamos a cabo, se muestra empíricamente que el mejor valor para el umbral es 1.

7.2.2 Selección del umbral

En una primera fase, hemos evaluado todos los descriptores con todas las configuraciones de división de la región de desgaste variando el número de parches necesarios para considerar la plaquita como dañada. En estos experimentos, nos hemos centrado en la métrica recall debido a la naturaleza de nuestro problema donde el impacto de una clasificación errónea de una plaquita dañada como utilizable es mayor que en el caso contrario.

La Figura 29 muestra el recall para todos los experimentos realizados usando todos los posibles valores de umbral. Como cabía esperar, en todos los casos, el umbral igual a uno ofrece el mejor recall independientemente del descriptor y de la configuración de las regiones de desgaste. A medida que aumenta el valor del umbral, el valor del recall disminuye. Sin embargo, no todas las configuraciones de las regiones de la zona de desgaste tienen el mismo comportamiento: los resultados usando FED, TBD y HED son muy dependientes del umbral. Por ejemplo, usando $ALBP_{16,2}$ y FED, el recall varía de 0.8 cuando el umbral es 1 a menos de 0.2 cuando el umbral es 4.

En contra a este comportamiento, HGD y SED son más invariantes al umbral utilizando ciertos descriptores. La diferencia con los mismos umbrales y descriptor que en el caso anterior es aproximadamente 0.1 usando HGD. Esto puede ser debido al alto número de divisiones que se extraen en estos dos últimos métodos: Cuanto más alto es el número de divisiones, más baja es la diferencia en términos de rendimiento entre los umbrales.

Además, es interesante destacar que el método que obtiene el mejor recall cuando el umbral es alto es $ALBP_{8,1}$ con una diferencia de más de un 20% con respecto al resto de métodos evaluados.

Por todo ello, hemos fijado el umbral a 1 para el resto de los experimentos.

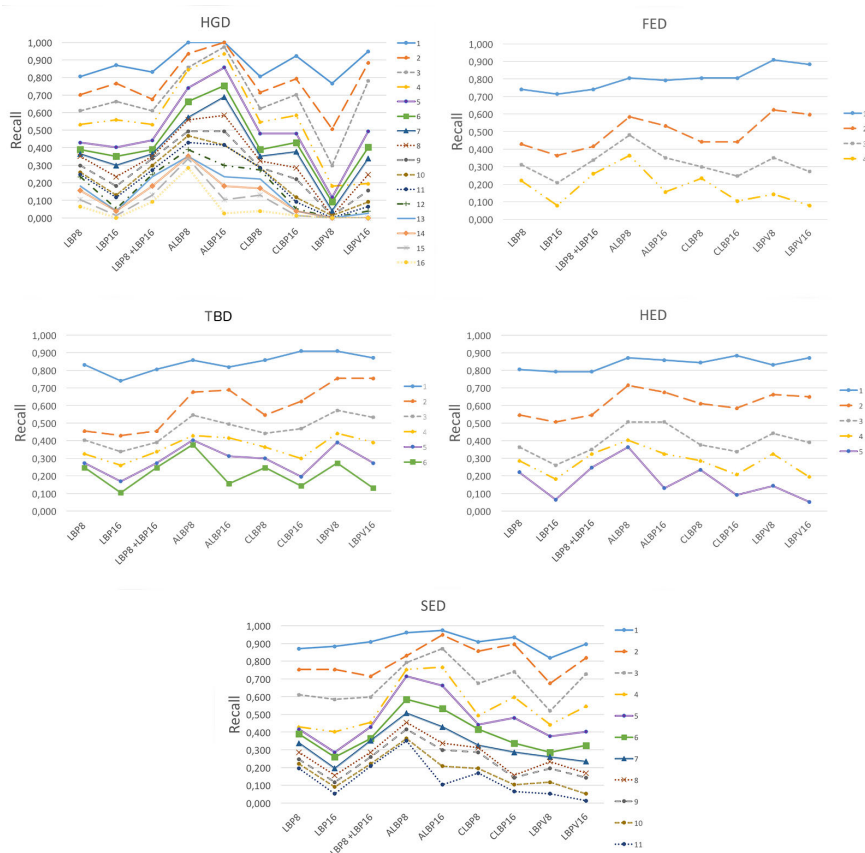


Figura 29: Recall para todas las configuraciones de la zona de desgaste usando diferentes descriptores y diferentes umbrales.

7.2.3 Configuración de la zona de desgaste y evaluación de la descripción

La Figura 30, muestra los resultados con umbral 1 y todas las posibles configuraciones de las regiones de desgaste y descripción. En este caso los resultados que se muestran tienen en cuenta cuatro diferentes métricas de evaluación: precisión, recall, accuracy y F-Score. Normalmente F-Score es la métrica más representativa debido a que tiene en cuenta los resultados de la precisión y el recall. Como podemos observar, el mejor FScore se obtiene al usar TBD y CLBP_{16,2} (0.909) y también al usar SED y LBP_{8,1}+LBP_{16,2} (0.903), donde + representa el operador de concatenación.

Otra consideración interesante es que aunque HGD y SED obtengan valores altos de recall con ALBP_{8,1} y ALBP_{16,2}, su precisión es demasiado baja como para poderse



Figura 30: Resultados de accuracy, precision, recall y F-Score de todos los métodos para todas las configuraciones de regiones y con umbral=1.

considerar una solución aceptable.

Teniendo en cuenta toda la información de las Figuras 29 y 30, podemos concluir que el mejor método es el formado por SED y $LBP_{8,1}+LBP_{16,2}$ debido a su alto recall (0.909), alto F-Score (0.903) y alto accuracy (90.26 %). Además, esta configuración es la más estable frente a cambios de umbral lo que permite a los expertos una mayor flexibilidad a la hora de realizar variaciones cuando la precision pudiera ser más importante de lo habitual.

7.3 Conclusiones

En conclusión, en este capítulo hemos propuesto un nuevo método de detección del estado de la herramienta de corte en procesos de torneado. Nuestro método propuesto es un método en tiempo real que utiliza dispositivos de bajo coste como Raspberry Pi o camaras monocromáticas como la Genie M1280.

Hemos presentado diferentes configuraciones para dividir el filo de corte de la plaquita de corte: HGD, FED, TBD, HED y SED..

Hemos demostrado que cuando al menos uno de los parches es clasificado como dañado, las probabilidades de que la plaquita entera lo esté aumentan considerablemente.

Usando una combinación de los descriptores $LBP_{8,1}$ y $LBP_{16,2}$ hemos logrado obtener un F-Score de 0.903 y una accuracy del 90.26% dividiendo la región de desgaste utilizando el método SED, lo que supera muchas de los sistemas desarrollados hasta la fecha. Sin embargo, la principal ventaja del método no es tanto su alto porcentaje de acierto sino el hecho de que no requiera segmentación de la zona de desgaste, cuyo coste computacional suele ser muy elevado y sujeto a errores, lo que permite al método realizarse sobre la propia herramienta sin perjudicar el rendimiento original.

8 Conclusiones de la tesis y trabajo futuro

En cumplimiento del punto 3º del artículo 19 del Reglamento de las enseñanzas oficiales de doctorado y del título de doctor de la Universidad de León, aprobado en Consejo de Gobierno el 25/9/2012, las conclusiones, contribuciones y trabajo futuro de esta tesis, han sido presentadas en el capítulo 9.